Personality and Social Psychology Review
2006, Vol. 10, No. 3, 186–213

How the Group Affects the Mind: A Cognitive Model of Idea Generation in Groups

Bernard A. Nijstad
Department of Psychology
University of Amsterdam

Wolfgang Stroebe
Department of Psychology
Utrecht University

A model called search for ideas in associative memory (SIAM) is proposed to account for various research findings in the area of group idea generation. The model assumes that idea generation is a repeated search for ideas in associative memory, which proceeds in 2 stages (knowledge activation and idea production), and is controlled through negative feedback loops and cognitive failures (trials in which no idea is generated). We show that (a) turn taking (production blocking) interferes with both stages of the process; (b) ideas suggested by others aid the activation of problem-relevant knowledge; and (c) cognitive failures are important determinants of brainstorming persistence, satisfaction, and enjoyment. Implications for group decision making and group recall are discussed.

Much of our working lives are spent in groups, and small group task performance has traditionally been an important topic for psychological research. Many studies have addressed the conditions under which groups are more versus less effective than individuals (see Kerr & Tindale, 2004, for a recent review). In general, this research has shown that groups can be effective, but that there are various problems associated with group work, such as motivation and coordination problems (Steiner, 1972), biased information sharing (Stasser & Titus, 1987), and an increased vulnerability to some cognitive biases and errors (Kerr, MacCoun, & Kramer, 1996).

One task that is often performed in groups is idea generation. Idea generation is one of the first stages of problem solving or decision making, in which potential solutions, decision alternatives, or hypotheses are generated (Osborn, 1953). Thus, teams of designers might generate ideas about new products (Sutton & Hargadon, 1996); teams of scientists might generate research hypotheses (Dunbar, 1995); or management teams might generate ideas on how to make their organization function more effectively (West & Anderson, 1996). Most people believe that idea generation is best performed in groups and presume that interaction with other people stimulates their creativity (e.g., Paulus, Dzindolet, Poletes, & Camacho, 1993). However, controlled research has consistently shown that people produce fewer ideas and ideas of lower quality when they work in a group as compared with when they work alone (Diehl & Stroebe, 1987; Mullen, Johnson, & Salas, 1991). Thus, contrary to popular belief, group interaction inhibits the ideation process.

The counterintuitive finding that groups are less effective idea generators than individuals has stimulated much research investigating the causes of this effect and ways to improve group idea generation. In this article, our aim is to review and integrate the research literature on group ideation (see for earlier reviews: Lamm & Trommsdorff, 1973; Paulus, Dugosh, Dzindolet, Coskun, & Putman, 2002; Stroebe & Diehl, 1994). Much of the early research has considered social–motivational factors, such as social comparison processes, evaluation apprehension, and free riding (e.g., Paulus & Dzindolet, 1993; Stroebe & Diehl, 1994). In contrast, we take a social–cognitive approach. We argue that idea generation is a cognitive task and that various effects of group interaction on performance can be interpreted as either *cognitive stimulation* or *cognitive interference* effects. Our social–cognitive approach entails specifying the cognitive processes underlying idea generation at the individual level and deriving hypotheses about how these individual level cognitive processes are affected by group interaction.

This research has been funded by Grant 575-31.007 of the Netherlands Organization for Scientific Research (NWO), awarded to W. Stroebe. We thank Carsten de Dreu for his help with the meta-analyses.

Correspondence should be sent to Bernard A. Nijstad, Department of Work and Organizational Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. E-mail: B.A.Nijstad@uva.nl

More specifically, we argue that there are strong parallels between idea generation and free recall. We use findings and theories from memory retrieval research, most notably Raaijmakers and Shiffrin's (1981) search of associative memory (SAM) model, to develop a model of idea generation that specifies how individual group members generate ideas. That model is called search for ideas in associative memory (SIAM).[1] We use this model to derive predictions on how "the group affects the mind" to cause cognitive stimulation and interference. Although some previous theorizing about group ideation has considered cognitive factors (Brown, Tumeo, Larey, & Paulus, 1998; Nagasundaram & Dennis, 1993), this is the first systematic application of an existing memory model to the group ideation process.

We first give an overview of the research findings in the field of group idea generation. We then describe the most comprehensive cognitive model of group ideation to date, the Brown et al. (1998) matrix model. We suggest that this model has some shortcomings and argue that it is useful to apply models of memory retrieval to group ideation. We next describe our SIAM model. Specific predictions are derived from SIAM and are contrasted with predictions derived from the matrix model. We discuss recent findings that are consistent with SIAM. In particular, we explain some of the more robust findings in the group idea generation literature and show that these have common cognitive underpinnings. In a final section, we evaluate our model.

Group Idea Generation: A Review of the Literature

Basic Findings

In the 1950s, Alex Osborn (1953), an advertising executive, suggested brainstorming as a method to improve the creativity of groups. To prevent premature evaluation of ideas and stimulate a more creative approach to problems, Osborn suggested that (a) emphasis during idea generation should be on quantity and not on quality of ideas, (b) participants should be encouraged to come up with unusual ideas, (c) they should combine and improve their ideas, and (d) criticism should be ruled out during idea generation, and evaluation of ideas should be deferred to a later stage of the problem solving process. Further, he argued that when these rules are applied "the average individual can think up twice as many ideas when working with a group than when working alone" (Osborn, 1957, p. 229).

The principles suggested by Osborn do indeed enhance idea generation. Thus, an emphasis on quantity has a positive effect on the number of ideas produced (Christensen, Guilford, & Wilson, 1957), and quantity and quality of ideas are strongly correlated (Stroebe & Diehl, 1994). Further, ruling out criticism has a positive effect on the average quality of ideas (Bartis, Szymanski, & Harkins, 1988), and brainstorming instructions in general enhance the generation of ideas (Parnes & Meadow, 1959). However, as mentioned earlier, the prediction that brainstorming is best performed in groups has not received support. Taylor, Berry, and Block (1958) performed the first test of Osborn's prediction. They compared four-person groups with so-called nominal groups, consisting of four individuals who worked alone. The nonredundant ideas of the nominal groups were pooled after the brainstorming session to arrive at a measure of productivity when group interaction neither facilitated nor inhibited productivity. They found, contrary to Osborn's prediction, that nominal groups generated many more ideas than interactive groups. Thus, interactive groups suffered a productivity loss, and the group process inhibited rather than stimulated idea generation. This result has since been replicated many times (see Diehl & Stroebe, 1987). Based on a meta-analysis, Mullen et al. (1991) could conclude that "productivity loss in brainstorming groups is highly significant, and of strong magnitude" (p. 18). The loss is relatively small for dyads but increases rapidly with group size.

The finding that group interaction inhibits idea generation is even more interesting because people generally presume it to be stimulating. Thus, most people believe group brainstorming to be more effective than individual brainstorming, which is true in the United States (Paulus et al., 1993), Europe (Stroebe, Diehl, & Abakoumkin, 1992), and Japan (Homma, Tajima, & Hayashi, 1995). Further, group members are more satisfied with their performance than individuals, although in fact they have generated fewer ideas (e.g., Nijstad, Stroebe, & Lodewijkx, 1999; Stroebe et al., 1992). Perhaps because of this illusion of group productivity (Paulus et al., 1993), group brainstorming still is widely applied in organizations.

Causes of the Productivity Loss

In the last decades, much has become clear about the causes of the productivity loss of brainstorming groups. There is some evidence for the role of *evaluation apprehension*. Despite the instruction not to criticize ideas, group members still seem somewhat anxious about sharing their ideas (Diehl & Stroebe, 1987), which is especially true for shy people (Camacho & Paulus, 1995). However, these effects are weak and do

[1]Goldstone (1994) has previously used the acronym SIAM in his similarity, interactive activation, and mapping model. We initially were unaware of that model. We had used SIAM in previous articles (Nijstad et al., 2002, 2003), so we did not change the name of our model.

not contribute much to the productivity loss of groups. *Free riding* or *social loafing*—the tendency to let other group members do the work because one cannot individually be held accountable for one's performance in a group—does not contribute much to the effect either: Making group members individually accountable for their performance does increase productivity, but this effect is relatively small (Diehl & Stroebe, 1987).

There is evidence for another motivational process, called *social matching*: High-performing group members have the tendency to match the rate of production of low performing members (Camacho & Paulus, 1995; Paulus & Dzindolet, 1993). High performers apparently do not want to compensate the lack of motivation or ability of others. Further, as a consequence of evaluation apprehension, free riding, and other factors, groups start out with a low rate of production. This may result in low performance standards, which are maintained in the rest of the session. If group members adhere to these low standards, they will perform at a similar low level (irrespective of ability). As a result of these processes, group productivity suffers (Paulus & Dzindolet, 1993).

Most support, however, has been found for the role of *production blocking* in the productivity loss of idea-generating groups. Production blocking refers to the fact that group members have to take turns when expressing their ideas. Research has shown that when people cannot express their ideas soon after they are generated because they have to wait for their turns, productivity strongly declines (Diehl & Stroebe, 1987, 1991). Obviously, individuals do not have this problem. Further, in larger groups production blocking is exacerbated, and therefore large groups suffer a larger productivity loss than small groups (and in dyads the loss is minimal).

Two points should be noted. First, overhearing the ideas of others is not necessary: Turn taking leads to productivity losses even without communication among group members (Diehl & Stroebe, 1987, 1991). The effect therefore is not due to distraction caused by overhearing others generate ideas. Second, at least for relatively small groups, the effect is not due to restrictions in speaking time. For group members, speaking time is shared, and group members may not have enough speaking time available to express all their ideas. However, Diehl and Stroebe (1991) restricted speaking time of individuals to only a part of the session while keeping session length and time to think constant (like members of four-person groups, individuals had 20 min to generate ideas, but could use only 5 min for speaking). This manipulation did not affect the productivity of individuals, and the productivity loss of groups was not reduced. This suggests that it is not overall speaking time that is important, but the blocking effect occurs when people cannot express their ideas at the time they choose.

If production blocking is the most important cause of the productivity loss of brainstorming groups, it is to be expected that eliminating production blocking will bring productivity of interactive groups to the level of nominal groups. This indeed is the case. It is possible to eliminate blocking by using procedures that do not require turn taking among group members. One of these is electronic brainstorming (EBS), in which people work at interconnected computers and can read one another's ideas on their screens. Because every group member can type simultaneously, production blocking no longer is a problem. Research has shown that as a result interactive EBS groups (with idea sharing) can perform at least at the level of nominal groups (Gallupe, Bastianutti, & Cooper, 1991). Further, when production blocking is introduced in electronic groups (e.g., by having group members take turns to type ideas), a productivity loss is found similar to the loss in verbal groups (Gallupe, Cooper, Grisé, & Bastianutti, 1994; Paulus, Larey, Putman, Leggett, & Roland, 1996).

Can Ideas of Others Also Be Stimulating?

Even more interesting is the finding that idea sharing can lead to productivity gains: Groups with idea sharing sometimes outperform nominal groups without idea sharing. EBS studies have shown that especially large groups (nine or more members) can achieve these gains (Dennis & Valacich, 1993; Valacich, Dennis, & Connolly, 1994). More recently, Dugosh, Paulus, Roland, and Yang (2000) have reported productivity gains in EBS groups with four members, but these gains were found only when group members were explicitly instructed to pay close attention to the ideas of others because they would be tested for recall of these ideas later on. This suggests that other studies did not find productivity gains in small EBS groups because participants did not pay attention to ideas of others. Paulus and Yang (2000) have similarly found productivity gains when using a writing method ("brainwriting"). They had four-person groups exchange their ideas on written notes in a round-robin fashion and found that groups with idea sharing outperformed those in which group members could not read one another's ideas.

Researchers have also exposed individual participants to preselected ideas or idea categories. Dugosh et al. (2000) had participants generate ideas while listening to an audiotape containing ideas generated by other people and found that listening to the tape enhanced performance. The stimulating effect, however, occurred only when participants paid close attention to the tape. Other researchers have exposed participants to category labels (e.g., for the topic "how can your university be improved," category labels such as "improve parking" and "improve facilities" were used). Coskun, Paulus, Brown, and Sherwood (2000) found (Experiment 1) that, in a 30 min session, offering 10 category labels once and simultaneously at the beginning of the session (simultaneous condition) was less effective than offering 1 new category every 3 min (sequential condition). The reason was that the less popular categories were explored more fully in the sequential condition, which enhanced performance especially in the later stages of the session. In Experiment 2, Coskun et al. found that offering only 2 categories was not effective, whereas offering 10 categories was (also see Dennis, Valacich, Connolly, & Wynne, 1996).

Diehl, Munkes, and Ziegler (2002) had individuals generate ideas at a computer. They manipulated independently whether participants could access category labels (11 categories) and whether they could access idea exemplars falling in that category, creating four conditions: (a) no stimulation control, (b) category stimulation only, (c) exemplar stimulation only, and (d) both category and exemplar stimulation. Participants could access category labels and/or ideas by clicking on buttons on their screens. Diehl et al. found that both the category labels and the idea exemplars were helpful, and more categories were surveyed in the three conditions with stimulation than in the control condition. However, access to both labels and exemplars had no extra stimulating value, and the number of categories surveyed was equal in the three experimental conditions. Further, exemplars (but not category labels) were detrimental to the generation of ideas within categories (fewer ideas per category). According to that study, therefore, stimulation will occur only when ideas of others give access to categories. However, idea exemplars at the same time inhibit the generation of ideas within categories.

The Illusion of Group Productivity

Researchers have also investigated why group members are more satisfied with their performance than individuals, although they have generated fewer ideas (the illusion of group productivity). Stroebe et al. (1992) argued that group members may be unable to distinguish between ideas they have generated themselves and ideas generated by other group members. This may lead to memory confusion and an overestimation of one's contribution to the group product, and thus to high levels of satisfaction. In line with the memory confusion explanation, Stroebe et al. have found that members of four-person groups claimed that no less than 60% of the ideas generated in a group session had also occurred to them. Further, they found that group members were much less accurate than individuals in identifying ideas that really were suggested by them.

Paulus et al. (1993) argued that social comparison processes are responsible for the illusion of group productivity. They presumed that group members compare their performance with that of other members and usually find that they have contributed similar numbers of ideas. Consequently, they are quite satisfied with their performance, although individuals who have brainstormed alone are less satisfied. In support of this explanation, Paulus et al. (1993) found that participants who worked individually but were provided with performance information of a fellow participant rated their performance more favorably than participants who did not receive this information. Clearly, social comparison processes contribute to the illusion of group productivity.

Conclusion

All in all, considerable progress has been made in the area of group idea generation. However, the processes that mediate these effects are still poorly understood. Thus, although production blocking has a large negative effect on performance, it is not clear why waiting for one's turn is so detrimental for idea productivity. For example, individuals may feel obliged to continue idea generation for most of their allotted time, whereas for group members speaking time is shared, so they can relax and let others do the talking (Diehl & Stroebe, 1987). Alternatively, ideas perhaps are forgotten while group members wait for their turn (Diehl & Stroebe, 1991). Similarly, access to ideas of others can stimulate creativity, but it is not clear when these stimulation effects will be found and why reading others' ideas can be stimulating.

Toward a Cognitive Model of Performance in Idea-Generating Groups

Our aim is to integrate and explain the most important phenomena of group idea generation, including the blocking effect, the stimulating effects of idea sharing, and the illusion of group productivity. To this end, we take a social–cognitive approach, which is laid out in this section. Before we turn to our SIAM model, which is at the core of our approach, we first discuss the most elaborate previous cognitive model of idea generation, the Brown et al. (1998) matrix model.

The Matrix Model of Idea Generation

Brown et al. (1998) conceptualized long-term memory (LTM) as a semantic network in their matrix model of idea generation. A semantic network is a collection of interconnected nodes representing various concepts (e.g., Anderson, 1995). The degree of association between nodes is represented as the strengths of the connections. Once a particular node is activated (becomes active in working memory), activation spreads to the connected nodes, and nodes that are strongly connected to the activated node have a higher probability of being activated (Collins & Loftus, 1975). People are assumed to sample ideas from this network and generate those ideas that are most strongly activated. After generating an idea, activation spreads to related ideas, and these ideas become activated and will be generated next.

Because the semantic network of a brainstormer is too complicated to model at the level of individual ideas, Brown et al. (1998) model brainstormers at the category level. They represent an individual brainstormer as a category transition matrix. First, the brainstormer generates an idea from one of *n* categories. Next, activation spreads to ideas connected to the generated idea. The matrix consists of *n* rows, representing the *n* different categories from which an idea is generated, and $n + 1$ columns, representing the following idea. The cells on the diagonal of the matrix contain the probability that an idea from a certain category is followed by an idea from the same category (a category repetition); the off-diagonal cells contain the probability that an idea is followed by an idea from a different category (a category change). The last $(n + 1)$ column gives the probability that no idea is generated in the next time interval (the "null category"). Initially, the probabilities on the diagonal are higher than the off-diagonal probabilities because semantically related ideas have stronger mutual ties. The probabilities in the null category are low at the beginning of the session but increase as more ideas are generated.

To deal with interactive brainstorming, Brown et al. (1998) made two further assumptions. To model production blocking, they assumed that when two or more "group members" generate an idea in a certain time interval, only one member can express this idea (who is determined at random). The other(s) forget their ideas when they cannot be expressed. Further, the idea generated at time $t + 1$ by person *i* no longer depends only on the idea generated by person *i* at time *t* but also on the idea expressed by a fellow group member at time *t*. The weight given to each of these is dependent on the amount of attention paid to the ideas of others (i.e., an attention parameter).

Brown et al. (1998) have run a number of simulations using the matrix model. A first result was that, because only one brainstormer can express his or her idea in each time interval, the number of spoken ideas per individual decreased with group size. However, the number of internally generated ideas per brainstormer increased with group size, but many ideas that could not be expressed were forgotten. Thus, according to this simulation, the blocking effect is due to idea forgetting. A second result was that the number of spoken ideas for an interactive pair increased with attention paid to the ideas of the partner. Thus, stimulation from ideas of others is possible, as long as group members pay enough attention to one another.

In a later version of the matrix model, Coskun et al. (2000) elaborated the model to include a working memory (WM) system. WM is conceptualized as "the set of currently active concepts" (p. 310), which might be different for each individual brainstormer. Generated ideas are added to WM, and if another group member speaks, that idea is added to WM as well. The generated idea at time $t + 1$, in this new version of the model, no longer depends only on the previously generated and/or expressed idea at time *t* but also on earlier ideas that still are in WM. However, in each simulation trial, the weights of the concepts in WM are reduced through a decay parameter. Earlier ideas thus have less influence on which idea is generated subsequently, whereas later ideas receive more weight. Using this new version of the model, Coskun et al. could effectively deal with the effects of exposure to category labels (which we described earlier).

The matrix model has several admirable features. The category transition matrix offers a relatively simple conceptualization of idea generation. The model can also deal with interactive brainstorming and may offer explanations for some of the phenomena of group brainstorming. However, there are two fundamental problems with the matrix model. First, the model has no rules to determine the values of the category transition probabilities. Thus, $n * (n + 1)$ transition probabilities have to be estimated from empirical data or arbitrary values have to be assigned. This issue takes on some importance because Brown et al. (1998) show in their simulations that the choice of specific probabilities can have a large effect on performance: "Convergent" brainstormers, who had high within-category transition probabilities (they were likely to stick to a category), were much less productive than "divergent" brainstormers, who had low within-category transition probabilities (they were more likely to switch categories). Thus, the model is sensitive to changes in category transition probabilities and has no theoretical rules to determine their values.

A second problem is that the matrix model is rather shallow on the cognitive processes underlying idea generation. It assumes that ideas can be retrieved from an associative network and that activation then spreads to connected ideas. However, ideas often are new solutions to a problem that cannot be directly retrieved from memory. Further, spreading of activation is an automatic process that lacks cognitive control, but within the brainstorming paradigm not every response is equally valid. Instead, ideas need to be relevant to the problem under consideration. Therefore, the generation of ideas must be controlled, a task which will require WM capacity. To conclude, although the matrix model has some promising features, it also has some shortcomings.

Free Recall and Idea Generation

Our aim is to develop a model that deals with these shortcomings. This model is inspired by the extensive empirical and theoretical work within the free recall paradigm. In this section, we go into the parallels between free recall and idea generation to justify our application of a memory retrieval model to group idea generation.

Parallels at the individual level. Recall of learned items differs from idea production because (most) ideas cannot be directly retrieved from memory but are new solutions that have to be generated. Nevertheless, theories and findings of free recall can provide insight in the process of idea generation. The reason is that ideas cannot be generated ex nihilo, but previously stored knowledge must be used to generate new ideas (cf. Amabile, 1983; Boden, 1990; Simonton, 2003). Thus, although idea generation is a process of production, it contains elements of retrieval: Previously stored knowledge must be activated to use it for the generation of ideas.

The free recall paradigm is in several ways similar to the brainstorming paradigm. In both cases, the number of responses is important (number of items recalled vs. number of ideas produced), and there is no fixed order in which responses must be given. However, the range of possible responses is restricted: In free recall only items that were presented during the learning trial are "valid;" in idea generation only ideas that are relevant to the problem are "valid." Of particular relevance to idea generation is the free recall of categorized words (i.e., the learned items can be classified in semantic categories). If we assume that ideas can be categorized as well, the reproduction of categorized lists might be similar to the production of ideas.

Several empirical phenomena associated with free recall have also been obtained in the brainstorming paradigm. First, in free recall the number of items recalled decreases over time: In every time interval fewer items are recalled than in the previous interval, until an asymptote is reached (e.g., Bousfield & Sedgewick, 1944). In brainstorming, performance also declines as the session proceeds, and this decline is more rapid in the early stages of the session (Diehl & Stroebe, 1991; Kanekar & Rosenbaum, 1972). Second, in categorized free recall, successive items are often from the same semantic category (e.g., Gruenewald & Lockhead, 1980). In brainstorming, a similar semantic clustering of ideas has been found (Diehl, 1991; Larey & Paulus, 1999). In free recall, these clusters are usually separated temporally as well: The time between items in two clusters on average is longer than the time between items within clusters. Although this effect has not been reported for brainstorming, it may occur there as well.

A phenomenon of particular interest in free recall is the part-list cueing effect (Slamecka, 1968). Participants learn a list of words and reproduce these words under two different conditions. In the control condition, participants simply recall as many words as they remember. In the experimental condition, a random subset of these words is offered as "retrieval cues." Surprisingly, participants who do not receive cues outperform those who do receive cues. However, Roediger (1974) has shown that when the number of categories in categorized free recall is high, retrieval cues can be helpful when cues give access to more categories than when no cues are given (and thus help to recover categories, not items). However, if cues provide more information than necessary to retrieve categories, recall is impaired. This phenomenon may be related to brainstorming because ideas offered by fellow group members may serve as retrieval cues for the other members. It has consequently been argued that the ideas of others should have effects similar to retrieval cues in the part-list cueing paradigm (Diehl, 1991; Diehl et al., 2002). Indeed, Diehl et al. found that exposure to ideas of others was helpful when they made more idea categories accessible, whereas it inhibited the production of ideas within categories.

Parallels at the group level. Several studies have investigated free recall in interactive groups. For example, Weldon and Bellinger (1997) compared the performance of three-person interactive and nominal groups on a free recall task. They found that interactive groups recalled fewer words (Experiment 1) and fewer propositional units from a story (Experiment 2) than did equivalent size nominal groups. They labeled this finding *collaborative inhibition* and drew parallels with the productivity loss in brainstorming groups.[2]

Basden, Basden, Bryner, and Thomas (1997) argued that collaborative inhibition is caused by a disruption of individual retrieval strategies in a group context. According to this explanation, each individual develops a retrieval strategy, which specifies the individually preferred organization of retrieval. When at the same time other group members retrieve information, the individual retrieval strategy is disrupted and is abandoned for a less effective strategy. This should explain why group recall is inferior to pooled individual recall.

[2]There is one important difference between these studies and the traditional brainstorming paradigm. In the group recall studies, there was no fixed time limit for recall, but participants continued until no new items were remembered for 30 sec. In brainstorming studies, participants usually get fixed time limits. There is evidence that the productivity loss of brainstorming groups is reduced when no time limits are given (Nijstad et al., 1999)—this is discussed in the Discussion section of the article.

In four experiments, Basden et al. (1997) found evidence for a disruption of individual retrieval strategies during group recall using lists of categorized words. In Experiment 1, they found that interactive three-person groups recalled fewer words than nominal groups (pooled recall of three individuals who worked alone). Moreover, they computed the adjusted ratio of clustering (ARC), a measure for the organization of recall (see Roenker, Thompson, & Brown, 1971). At maximum clustering, the ARC is 1.00, and an ARC of 0 implies chance levels of clustering. Basden et al. found that clustering was higher for individuals (M = .66) than for group members (M = .24). Collaboration thus interfered with the individual organization of recall. Interestingly, Diehl (1991) also found that clustering in idea generation was lower for group members than for individuals. This finding was replicated in Experiment 2.

In Experiment 3, Basden et al. (1997) eliminated the performance difference between interactive and nominal groups. They had interactive and nominal groups recall items from six categories. Half of the participants were instructed to only recall items from two of the six categories, leaving the other four categories to the other two group members. In this situation, interactive groups recalled as many items as did nominal groups. However, the collaborative inhibition effect was present when all group members recalled items from all six categories. Apparently, participants' retrieval strategies were not disrupted when each group member recalled items from different categories. Finally, in Experiment 4, they had participants recall items category by category. Thus, they first had to exhaust one category before switching to the next. This should reduce collaborative inhibition because there should be no switching between categories and thus no disruption of retrieval strategies. Results were in line with these predictions.

To conclude, group recall is less effective than pooled individual recall, a finding that is similar to the productivity loss found in brainstorming groups. Further, the effect appears to be due to a disruption in the organization of recall of individual group members. However, at this point, one thing remains unclear. In these studies, members of interactive groups had to take turns to recall items while at the same time they could overhear each other. Collaborative inhibition might thus either result from production blocking (waiting for one's turn) or from overhearing other participants. We return to this in the Discussion section of the article.

SIAM

Because of the parallels between free recall and idea generation, models of memory retrieval might be fruitfully applied to better understand (group) idea generation. In this section, we use Raaijmakers and Shiffrin's (1981) SAM model of memory retrieval (also see Raaijmakers, 1993) to develop our SIAM model of idea generation. As an example to clarify different constructs, we use a brainstorming session about the topic "how can the number of tourists visiting your city be increased."

Basic assumptions. SIAM assumes two memory systems: LTM and WM. LTM is assumed to be essentially permanent and has unlimited capacity. It is a richly interconnected network, with numerous levels, categories, and associations. Like SAM, SIAM assumes that LTM is partitioned into *images* (no visual or spatial representation is implied). Images are knowledge structures that consist of a central concept and a number of features of that concept or associations with that concept. For example, the image "hotel" has features such as "has rooms," "has a lobby," and "is expensive." Images have fuzzy boundaries, may overlap to a considerable degree, and have mutual associations. The image of "hotel" may, for example, be associated with the image "restaurant."

WM has limited capacity and functions as a temporary storage system. Conscious operations, such as rehearsal, recognition, and decision making take place in WM. Following Baddeley and Hitch (1974; Baddeley, 1986, 1990), we assume three subsystems of WM: the central executive, the phonological loop (to temporarily store auditory and verbal information), and the visuo-spatial sketchpad (to temporarily store visual information). The central executive is involved in tasks that require attention and in the activation of material that has previously been stored in LTM (Baddeley, 1996; also Cinan, 2003). The other two systems are assumed to be "slave systems" of the central executive. The three subsystems are limited in capacity.

Like SAM, SIAM assumes a repeated search process in associative LTM. The most important assumption is that this search process proceeds in two stages. Because ideas cannot be generated without reference to prior knowledge, the first stage is the activation of knowledge in LTM. To activate knowledge means that images from LTM are temporarily stored in WM (WM may also be conceptualized as the active part of LTM; e.g., Baddeley, 1996). It is assumed that only one image may be active (in WM) at the same time (cf. Cantor & Engle, 1993; Daneman & Carpenter, 1980). In the second stage, the features of the image are used to generate ideas by combining knowledge, by forming new associations, or by applying knowledge to a new domain (cf. Mednick, 1962; Simonton, 2003). For example, one might generate the idea of making hotels less expensive by combining one's knowledge about features of the image "hotel" ("is expensive") with knowledge about tourists ("like cheap accommodation").

Similar to SAM, SIAM assumes that knowledge activation is cue dependent and probabilistic. A search cue is generated in WM, and this cue is used to activate images in LTM. In SAM, contextual cues (cues present during learning) and previously recalled items are used as cues. SIAM assumes that the problem definition is an important cue because activated knowledge must be relevant to the brainstorming problem. Previously generated ideas can also be added to the search cue, and other cues may be generated through "event sampling" (thinking about one's own previous experiences, e.g., about one's experiences as a tourist). Which image is sampled depends on the associations between the search cue and the features of the image. The accessibility of images thus depends on which search cue is being used (cf. Tulving & Pearlstone, 1966). The sampling of images is presumed to be sampling with replacement.

SIAM further assumes that when an image is activated it will be more strongly associated with the elements of the search cue, which makes it more likely that the same image is sampled again. Similarly, after an idea is generated, it will be more strongly associated with the image from which it is generated and with the elements of the search cue. Following SAM, this is called incrementing. Finally, similar to SAM, we assume that the process of idea generation is controlled through negative feedback loops and cognitive failures. People monitor their cognitive processes and keep track of instances in which they were unable to activate (new and relevant) knowledge or generate (new) ideas. These failures serve as input in a negative feedback loop.

A description of SIAM. A conceptual flowchart of SIAM is presented in Figure 1 (cf. Raaijmakers & Shiffrin, 1981, Figure 2 and Figure 3, pp. 97–98). According to the model, the production of ideas is a controlled, associative process that proceeds in two stages. The two stages are represented by two feedback loops: the *image retrieval loop* and the *idea production loop*. In the first stage information is retrieved from LTM, which is cue dependent and probabilistic. Because knowledge must be relevant to the problem, the problem definition is used in the search cue. Further elements of a search cue may be obtained through event sampling, as described previously. Because the retrieval of information is cue dependent, and a search cue needs to be generated, it is presumed to be a relatively effortful process involving the central executive component of WM.

The result of this search is the activation of a localized set of information (i.e., image). When an image has been retrieved from LTM, the features and associations of the image become accessible. When, for example, thinking about attracting tourists to Amsterdam, the image "canal" might be activated with its features, such as "has bridges" and "has boats." These features must be used to generate ideas. This is what is done in the second stage of the process, the idea production loop. It is assumed that one image has several features and these features can be combined with one another or with elements of the search cue in different ways to produce one or more ideas. For example, the feature "has bridges" may lead to the idea of decorating these bridges. When a participant has generated an idea, this idea must be temporarily stored in WM, after which it can be articulated. The participant will continue to generate ideas from the same image as long as he or she is able to do so. Using one image, successive ideas can be generated relatively quickly because it is not necessary to activate new knowledge; these ideas will be semantically related to one another. The generation of ideas using the features of the image is presumed to be relatively automatic (i.e., does not require much capacity of the central executive component of WM). In this way, a train of thought—a rapid and relatively automatic accumulation of semantically related ideas—may arise.

When a number of ideas have been generated, it will become harder to generate new ideas using the same image because it will be more likely that an idea that has already been mentioned is generated again. This likelihood increases further because of incrementing: When an idea has been generated, it is assumed to be more strongly associated with the problem, the search cue, and the image from which it has been generated. This will increase the chance that a particular idea is generated again (cf. fixation effects, e.g., Smith, 2003).[3] Generating the same idea again or failing to generate an idea will be called a failure. These failures are monitored and serve as input in a negative feedback loop. When the number of failures exceeds a stop criterion, the participant will leave the idea production loop. He or she will try to activate a new image using a new search cue. In this new cue, one or more recently generated ideas may be incorporated. However, in some cases the participant may also fail to activate a new image. These failures are monitored as well. When a stop criterion in the image retrieval loop is reached (i.e., when a number of successive searches are unsuccessful), the process will be terminated because it is becoming more difficult to generate additional ideas, and the impression arises that not many additional ideas can be generated. This negative feedback procedure is similar to the one employed by SAM.

[3]The generation of "old" ideas might be one reason why productivity declines in the course of a session: Old ideas are strongly connected to the problem and have a high probability of being generated again. As with fixation effects in problem solving (Smith & Blankenship, 1989) and recall (Smith & Vela, 1991), time off the task ("incubation") might resolve these problems because during off-task time activation of earlier ideas may decline. Indeed, there is some evidence that short breaks during brainstorming enhance performance considerably and prevent the usual sharp decline of performance over time (Paulus et al., 2002).

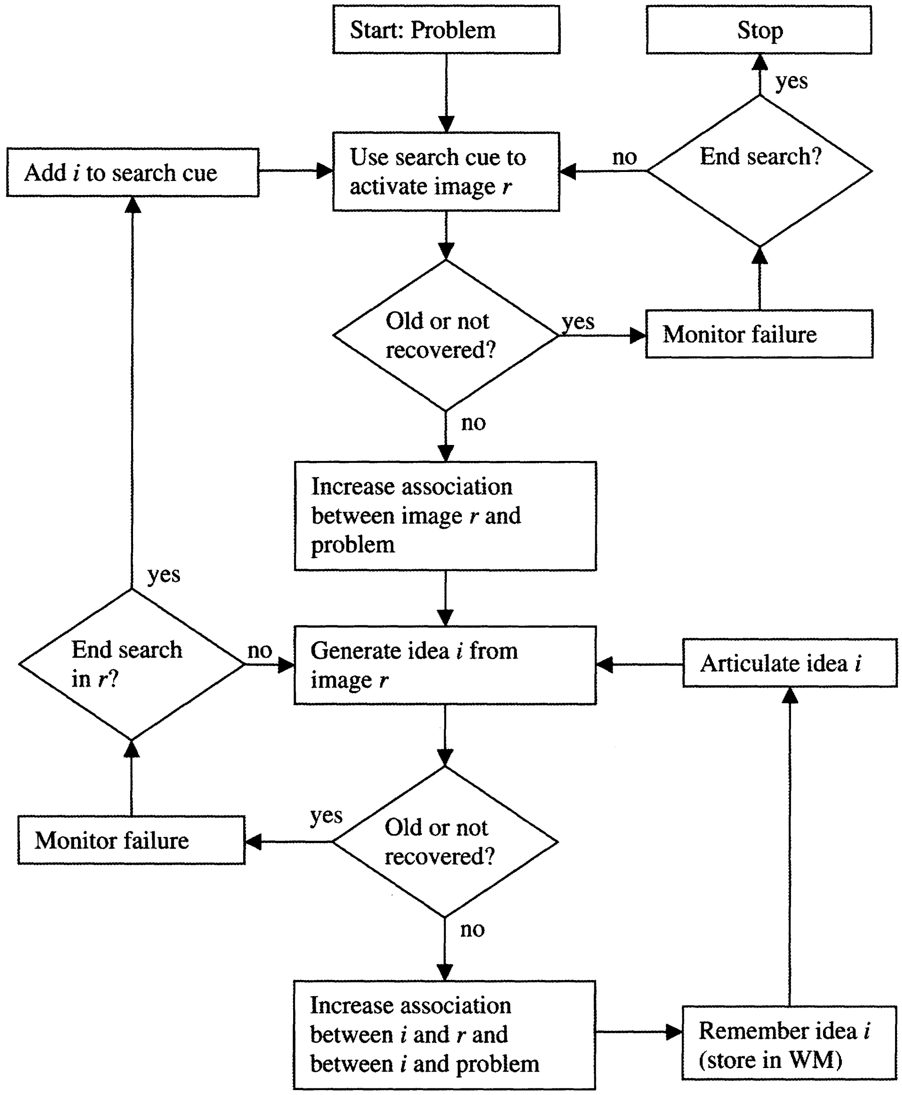

Figure 1. Flowchart of SIAM.

SIAM versus the matrix model. SIAM and the matrix model are similar in the sense that associative processes are assumed to underlie idea generation. There are two important differences. First, whereas the matrix model emphasizes the structural properties of associative networks through the category transition matrix, SIAM emphasizes active search processes. The idea generation process in SIAM is more active, controlled, and effortful because knowledge activation is assumed to rely on search cues and the central executive component of WM. The matrix model, in contrast, assumes passive and automatic spreading of activation. As will be clear, this leads to divergent predictions on a number of issues. Second, the search process in SIAM is controlled through cognitive failures and negative feedback. Although cognitive failures can be modeled through the null category of the matrix model, Brown et al. (1998) do not theorize about the role of failures.

In the following sections, we use SIAM to generate specific and testable predictions about idea generation at the individual level and the group level.[4] We argue that some of these predictions would not be made by the Brown et al. (1998) matrix model. We also describe recent evidence that is consistent with SIAM. First, we describe results pertaining to individual idea generation. Then, we

[4]We should note two things. First, we describe a research program of almost a decade, in which findings have sometimes led us to revise our theory. What we present here is the current state of the theory. Second, we have not quantified SIAM in a simulation model. Because there are no a priori limitations to the number of ideas that are possible for a brainstorming topic, it is impossible to establish the strength of interidea associations. In SAM, the strengths of the associations between items are based on presentation time during learning. Because in idea generation there is no prior learning trial, a similar procedure is not possible. The matrix model partly solves this issue by modeling intercategory associations rather than interidea associations. However, the model has no rules to specify the category transition probabilities. It is unclear how to solve this problem, and we chose not to quantify SIAM.

describe how group interaction affects these individual level processes.

Testing SIAM I: Individual Idea Generation

SIAM is a model of individual level idea generation. Although we are eventually interested in the explanation of group level phenomena, it is important to first establish whether SIAM's predictions are confirmed at the individual level. In this section, we test two classes of predictions: clustering in idea generation, and the role of failures.

Semantic and Temporal Clustering

Methodological considerations. SIAM assumes that that one image can be used to produce different ideas employing the features of that image. The resulting ideas will be semantically related and can be generated relatively quickly because no new image needs to be activated. Unfortunately, it is not possible to directly observe the activation of images. However, it is possible to code ideas into semantic categories using a predefined category system. We assume that when two successively generated ideas are coded in the same category, they are generated from the same image. Although this operationalization is imperfect (ideas might be generated from the same image but may be coded in different categories, or vice versa), this would generally only lead to an increase in measurement error and a conservative test of our hypotheses.

We have used two category systems (for two different topics) that were developed by Diehl (1991). In these category systems, a number of subgoals are crossed with a number of means to achieve these goals, resulting in a goal-by-means matrix of semantic categories. One category system is concerned with the environment ("what can people do to preserve the environment"). It consists of 10 goals (e.g., "reduce water use or pollution," "protect animals and plants") and five means (e.g., "consumption," "organization and action"), resulting in 50 categories. When categorized, it can be established whether an idea is a category repetition (it is coded in the same category as the previous idea) or change (it is coded in a different category). Presumably, category repetitions reflect that ideas were generated from the same image. Further, one can compute a number of additional measures, such as the ARC (Roenker et al., 1971). Figure 2 illustrates the different measures we use in this article.

Hypotheses. Our first hypothesis is that successively generated ideas often are semantically related (i.e., when they are generated from the same image). Thus, there should be semantic clustering in idea production, and an idea from one semantic category should more often be followed by an idea from the same category (a category repetition) than would be expected according to chance (H1). The matrix model would also make this prediction because it assumes that semantically related ideas have relatively strong mutual ties, and that therefore category repetitions are more likely than category changes.

A second prediction is that on average it takes less time to generate semantically related ideas than it takes

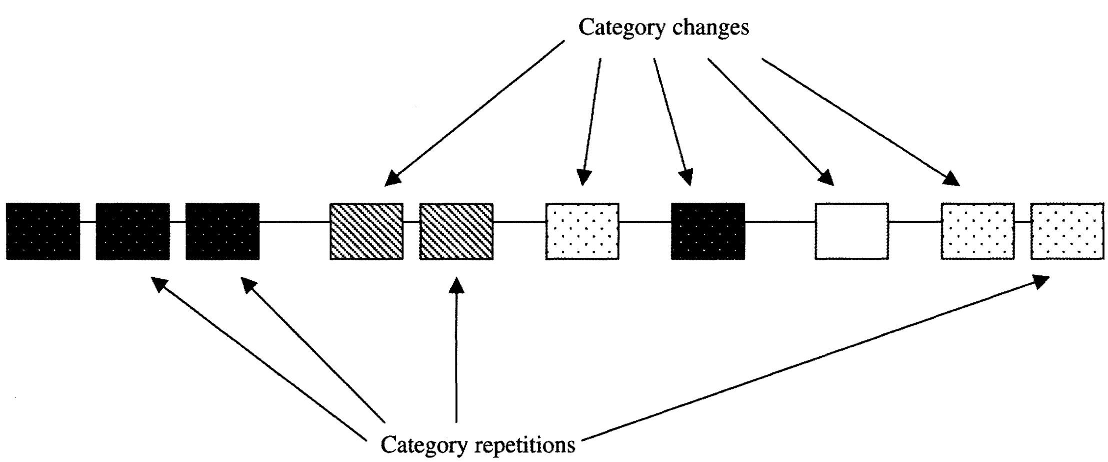

Figure 2. A hypothetical sequence of ideas. In this hypothetical sequence of ideas, each box represents an idea, whereas each different pattern represents a different semantic category. In this example, 10 ideas have been generated. This is our measure of productivity. Four of these ideas represent category repetitions because they are coded in the same category as the previous idea. There are six semantic clusters, separated from each other by category changes (an idea is coded in a different category than the previous one). To illustrate our finding that category repetitions are faster than changes, we used shorter distances between boxes for category repetitions than for category changes. The average cluster length in this case is 10 ideas divided by 6 clusters = 1.67. In total, four different categories have been used in the example, and this is our measure of idea diversity (number of categories sampled at least once). Dividing the number of ideas by the number of categories results in a within-category fluency (the average number of ideas per category, 2.5 in this case). Finally, the Adjusted Ratio of Clustering (ARC) can be computed (see Roenker et al., 1971, for the formula), and the ARC is .50 in this case.

to generate semantically unrelated ideas. Before a category change (drawing an idea from a different category than the previous idea), a new search cue has to be developed, and a new search of memory is necessary. This will take some time and makes category changes slower than category repetitions (H2). The matrix model would probably not make this prediction (although it is not explicit on this) because it assumes that ideas are directly activated through spreading of activation. When activation directly spreads to ideas in another category (or to ideas in the same category), there is no reason to expect response latency differences for category repetitions and changes.

Our reasoning further implies that it is more efficient to have high levels of clustering (many category repetitions) and to switch categories only if necessary. Especially returning to a previously used category implies more category changes than "depleting" the category in one go. Because category changes are slower, more switching between categories will lead to fewer ideas being generated within the same time interval. Thus, everything else being equal, clustering should be positively correlated with production (e.g., the number of ideas; H3). The matrix model would not make this last prediction but would even make the opposite prediction, that clustering is negatively associated with productivity (also see Larey & Paulus, 1999). In the matrix model, higher levels of clustering can only be modeled with relatively high within-category transition probabilities. Because probabilities must add to one, the between-category transition probabilities necessarily are lower, which makes category switches less likely. As a result, some categories become less accessible, and fewer ideas are generated in those categories. Indeed, Brown et al.'s (1998) simulations show that relatively high within-category transition probabilities lead to fewer category changes (i.e., higher clustering) and to fewer ideas than relatively low within-category transition probabilities. This suggests that, according to the matrix model, clustering is negatively associated with productivity (although Brown et al. do not explicitly make this prediction).

Results. Table 1 lists the results from four different experiments, taken from Nijstad, Stroebe, and Lodewijkx (2002, 2003). In these studies, individuals worked at a computer terminal. They were asked to generate as many ideas as possible in a 20-min session. We used topics such as the previously mentioned environment problem and the health problem ("what can everybody do to improve or maintain one's own health"). Table 1 shows only the data from our control group participants—the other results are described later—and these participants simply generated ideas without further instructions or additional manipulations. As can be seen, the different studies resulted in very similar findings. First, in all studies the ARC was positive and significantly different from zero, ranging from .20 to .41 (chance clustering would imply an ARC of zero).[5] Thus, there is evidence for semantic organization in idea production, confirming H1. Second, response latencies of category repetitions were shorter than those of category changes. This was true in all studies, the difference ranging from about 6 sec to 12 sec. This is consistent with H2. Third, in all studies, the ARC was positively correlated with production (the number of ideas generated), and correlations were moderate to high, ranging from .41 to .71, which suggests that H3 is confirmed as well. It should be noted that this is not an artifact because the ARC is mathematically independent of the number of ideas generated. Thus, our data supports SIAM's predictions, and there is semantic as well as temporal clustering in idea generation. Further, higher levels of clustering are associated with higher productivity (more ideas).

For these effects, we also conducted meta-analyses using the Hedges and Olkin (1985) approach. We used DSTAT 1.11 (Johnson, 1989) to compute Cohen's *d*, and a 95% confidence interval around *d*. We also computed the homogeneity statistic Q_w. A significant Q_w indicates that effect sizes differ across observations. With an effect size of .80 or higher considered large (e.g., Rosenthal, 1995), all our effect sizes were large (see Table 1). Moreover, none of the confidence intervals contained zero, and none of the Q_w statistics were significant (all $Q_w < 5.00$, $p > .15$). Thus, the four observations can be considered independent replications of the same effects, and we can be confident that these effects are not due to chance.

Cognitive Failures

Methodological considerations. According to SIAM, cognitive failures are monitored throughout the brainstorming session and play an important role. We define failures as a failure to generate a new idea or activate a new image while trying (see Figure 1). Failures may consist of either generating no idea at all (or failing to activate a relevant image) or generating an idea (or activate an image) that has already been mentioned (activated). In our studies, we have measured cognitive failures in our postexperimental questionnaires. This was done with three items: "how difficult was it to keep on generating ideas," "how often were you unable to generate ideas," and "how often did an idea you previously generated occur to you again." In all our studies, these items were positively intercorrelated and a com-

[5]These levels of clustering are not very high when compared with clustering in free recall (Basden et al., 1997, found levels of clustering of .60–.70). However, categories used in free recall are semantically more distinct (e.g., boy's names vs. animals) than are the ideas we categorize. With shorter "semantic distance" between categories, category changes are more likely.

Table 1. *Summary of Data Related to Clustering in Individual Idea Generation (Control Conditions Only)*

	Study				
Measure/Parameter	**Nijstad, Stroebe, and Lodewijkx, 2003, Exp. 1 (N = 10)**	**Nijstad et al., 2003, Exp. 2 (N = 20)**	**Nijstad et al., 2003, Exp. 3 (N = 17)**	**Nijstad, Stroebe, and Lodewijkx, 2002 (N = 14)**	**Effect Size d (Confidence Interval)**
ARC: M (SD)	.41 (.18)	.20 (.15)	.29 (.10)	.30 (.14)	—
ARC > 0	$t = 7.13, p < .001$	$t = 6.25, p < .001$	$t = 11.29, p < .001$	$t = 7.69, p < .001$	1.87 (1.44, 2.30)
Response Latencies (s)					
Category Changes	26.13 (7.66)	25.37 (14.86)	30.02 (8.20)	41.53 (12.06)	—
Category Repetitions	19.89 (6.38)	19.52 (12.54)	21.64 (9.56)	29.33 (8.78)	—
Difference	$t = 4.48, p < .01$	$t = 2.09, p = .05$	$t = 4.20, p = .001$	$t = 5.22, p < .001$	0.91 (0.53, 1.28)
Correlation ARC—Productivity	$r = .59, p = .07$	$r = .49, p = .03$	$r = .71, p = .001$	$r = .41, p = .15$	1.26 (0.87, 1.65)

Note: ARC = adjusted ratio of clustering, d = Cohen's d for effect sizes.

posite scale yielded internal consistencies (Cronbach's α) in the range of .61 to .73.

We deliberately measured failures subjectively, whereas one could also measure failures in a more objective way. For example, one could measure how often a long silence occurred in a brainstorming session and argue that longer silences imply failures. However, during a silence participants may not have been actively trying to generate ideas (e.g., in a group, a person may listen to another member). According to our definition of failures, it is essential that someone was trying to generate ideas but failed. Nevertheless, our subjective measure of failures should correlate with more objective measures. This indeed is the case. We computed the number of ideas produced in the last 4 min of the session and correlated this with the item "how difficult was it to keep on generating ideas." We found correlations of –.39 and –.35 in two different studies. Further, in four studies we computed the number of ideas with long response latencies (latencies of more than one standard deviation above average response latency in that study) and correlated this with the item "how often were you unable to generate ideas." The correlations were .35, .38, .41, and .46. These results show that our subjective measure is valid.

Hypotheses. We hypothesize that several aspects of subjective task experience and performance are related to the (subjectively experienced) number of cognitive failures. First, a brainstorming session will be less enjoyable when a high number of searches result in cognitive failures because experiencing large numbers of failures will be frustrating (H4). Second, when high numbers of failures occur, the individual may think that he or she is not performing well. Thus, failures should be negatively related to satisfaction (H5). Third, throughout the session the number of failures should rise, partly because more "old" ideas will be generated (i.e., ideas that have already been mentioned). This rise in the number of failures will eventually lead to the impression that not many ideas are still possible and thus to low expectancy (i.e., the belief that investing additional effort will not lead to many more ideas; cf. expectancy value theory, Vroom, 1964). Thus, expectancy will be negatively related to failures (H6). Finally, low expectancy (because of many successive failures) will, in turn, lead to a tendency to end the brainstorming task (H7; also see Figure 1).

Results. Table 2 lists the correlations between our composite failures scale and three measures of subjective task experience: task enjoyment ("how much did you enjoy the brainstorming task"), expectancy ("how many ideas will you still be able to generate"), and satisfaction ("how satisfied are you with your individual performance"). The first four columns represent the same studies as in Table 1, and the last two columns are derived from Nijstad et al. (1999). In the latter studies, participants did not type their ideas but articulated them aloud, and only those participants who worked individually are included in the table (other results are described following). As can be seen, correlations were weakest (but consistent) for task enjoyment (–.18 to –.33), stronger for expectancy (–.18 to –.49), and strongest for satisfaction (–.28 to –.59).[6]

We also conducted meta-analyses on these correlations. Table 2 reports Cohen's d, and the 95% confidence interval around d. In all cases, d was moderate to high, and the confidence intervals did not include zero. We also computed rho (ρ), which is the average correlation (corrected for sample size) between the variables. For the correlation between failures and enjoyment, rho was –.23, and the 95% confidence interval was –.34 to –.12. For expectancy, rho was –.39 (–.65, –.13), and for satisfaction it was –.40 (–.71, –.09). The homogeneity statistic (i.e., Q_w) was not significant for enjoyment and expectancy ($p > .10$), but it was significant for satisfaction, $Q_w = 11.20, p = .05$. The large correlation between failures and satisfaction in Nijstad et al. (2003) Experiment 1 ($r = -.59$) was responsible for this, and when this correlation was removed from the set, Q_w was no longer significant, $Q_w = 3.79, p = .43$. For the five remaining studies, rho was –.36, and the 95% confidence interval was between –.59 and –.13. It can be concluded that failures are associated with lower task enjoyment, expectancy, and satisfaction, and H4, H5, and H6 are confirmed. It should be noted, though, that it is not possible to establish a causal relation from these correlations.

Further evidence for the role of failures comes from the Nijstad et al. (1999) studies. Participants in those studies were not given a time limit but could continue brainstorming until they felt it was a good time to stop. After they had stopped, we asked them to describe their reasons for stopping. There may be different reasons for ending the session, including "I was getting bored" (e.g., Martin, Ward, Achee, & Wyer, 1993) or "I was satisfied with my performance" (cf., Paulus & Dzindolet, 1993). However, the reason that was mentioned most often (84.6% in Experiment 1; 81.3% in Experiment 2) was "I was running out of ideas." After some time it gets more difficult to generate ideas, and the number of failures increases. This leads to low expectations to be able to generate additional ideas (low expectancy). This is a powerful motive to end the task: When asked to indicate how many ideas they thought

[6]The intercorrelations among these items were weak to moderate, indicating that they do not represent a common underlying construct. In general, satisfaction and enjoyment were positively correlated (average $r = .31$). Enjoyment was not strongly correlated with expectancy (average $r = .15$), and expectancy and satisfaction were not strongly correlated either (average $r = .14$). Further, the correlations presented in Table 2 do not exceed .60, indicating that failures (empirically) are a distinct construct as well.

Table 2. *Correlations Between Failures and Enjoyment, Expectancy, and Satisfaction in Six Studies (Individual Idea Generation Only)*

	Study						
Measure	**Nijstad, Stroebe, and Lodewijkx, 2003, Exp. 1 ($N = 50$)**	**Nijstad et al., 2003, Exp. 2 ($N = 46$)**	**Nijstad et al., 2003, Exp. 3 ($N = 49$)**	**Nijstad, Stroebe, and Lodewijkx, 2002 ($N = 63$)**	**Nijstad, Stroebe, and Lodewijkx, 1999, Exp. 1 ($N = 26$)**	**Nijstad et al., 1999, Exp. 2 ($N = 32$)**	**Effect Size *d* (Confidence Interval)**
Enjoyment	–.21	–.18	–.23	–.27*	–.21	–.33†	–.48 (–.65, –.31)
Expectancy	–.49***	–.48**	–.18	–.38**	–.44*	–.47**	–.86 (–1.04, –.68)
Satisfaction	–.59***	–.38**	–.28*	–.31*	–.46*	–.50**	–.88 (–1.06, –.70)

Note: d = Cohen's d for effect sizes.
†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

they would still be able to generate on a 9-point scale ranging from 1 (*none at all*) to 9 (*very many*), participants answered that they would not be able to generate many additional ideas (Experiment 1: $M = 1.96$; Experiment 2: $M = 2.34$). As can be seen in Table 2, expectancy was correlated with failures. Although these results are correlational in nature, they do suggest that failures may play an important role in the process of idea generation and in the decision to end the session, as was predicted in H7.

Conclusion

In conclusion, our predictions regarding individual idea generation are supported. First, there appears to be semantic and temporal clustering in idea generation: Semantic clustering is higher than chance, and category repetitions are quicker than category changes. As a consequence, clustering is positively correlated with productivity (number of produced ideas). We noted that the Brown et al. (1998) matrix model would not predict some of these effects. Second, cognitive failures (instances in which no new idea is generated) are negatively correlated with task enjoyment, satisfaction, and expectancy. Failures also appear to play a role in the decision to end a brainstorming session: When many failures occur, people have lower expectancy that going on will lead to more ideas and therefore decide to end the session. The Brown et al. matrix model is silent about the role of failures. We thus conclude that SIAM receives support at the individual level.

Testing SIAM II: Effects of Communication on Cognition

SIAM is a model of individual level cognitive processes. We propose that these individual level processes are strongly affected by group interaction. Using relatively simple assumptions, SIAM can be used to derive specific predictions on the way group interaction affects cognitive processes of group members. In this section, we describe three ways in which interaction affects cognitive processes and how this process of mutual influence relates to group effectiveness. We discuss the effects of mutual production blocking, the effects of exposure to others' ideas, and the effects of communication on cognitive failures. First, however, we go into some additional assumptions that are necessary to apply SIAM to individuals who work in groups.

Additional Assumptions

Two things change as a result of group interaction. First, group members have to take turns to express their ideas (i.e., production blocking). This implies that not every idea can be immediately expressed, but that delays arise between the generation and expression of ideas. Second, participants overhear one another's ideas. For the effects of these two factors on cognitive processes of group members, we make the following four assumptions:

1. Short-term forgetting: When a participant has generated an idea but cannot express it immediately (because someone else is talking), the idea needs to be stored in WM. Based on findings from delayed recall (e.g., Haarmann & Usher, 2001; Loftus, 1974), we assume that the longer an idea has to be stored, the more likely it becomes that the idea is forgotten.
2. Search inertia: When a participant is remembering an idea, to express it later, he or she cannot activate a new image because of limitations in WM. Thus, either the idea is forgotten or no new search of memory is possible.
3. Delay monitoring: The participant needs to monitor delays to seize the opportunity to express ideas when they occur. This task requires attention and involves the central executive component of WM.
4. External cues: Ideas of others serve as external cues and, provided that they are attended to, are automatically added to the search cue to activate images in WM (cf. SAM's modeling of experimenter provided cues in the part-list cueing paradigm; Raaijmakers & Shiffrin, 1981).

Production Blocking: The Cognitive Interference Hypothesis

Background. Production blocking (i.e., turn taking among group members) is an unintended yet inevitable consequence of communication in verbally interacting groups. Production blocking seems to be the major (but not the only) cause of the productivity loss of groups. The effect does not seem to be due to distraction caused by overhearing others' ideas because it also occurs in the absence of communication among group members (Diehl & Stroebe, 1987, 1991). The blocking effect also does not appear to be due to insufficient time to express ideas in a group setting but rather occurs when ideas cannot be expressed soon after they have been generated (Diehl & Stroebe, 1991). This raises the question of what does cause the blocking effect.

There are several indications that the blocking effect is due to cognitive interference. Diehl and Stroebe (1991) tried to make turn taking more predictable. Participants could sign up on a speakers' list so they would know when they could express their ideas. However, a combination of this list and communication (through headphones) led to fewer ideas, suggesting cognitive overload when attention had to be divided among idea generation, listening, and the speakers' list. In another study, Diehl and Stroebe (1991) provided participants with notepads so they could write down their ideas to not to forget them while they were waiting for their turn. Although a notepad slightly improved performance in the absence of communication, it impaired performance when communication was possible, again suggesting cognitive overload. These findings are consistent with a cognitive interference hypothesis of production blocking.

Methodological considerations. Previous studies of the blocking effect have looked only at effects of blocking on productivity (the number of ideas generated). In addition to productivity, we look at five other dependent variables, based on the coding of ideas in semantic categories: (a) cluster length (average number of successive ideas coded in the same category), (b) number of clusters (number of times a new semantic cluster was started), (c) diversity (number of categories surveyed at least once), (d) within-category fluency (average number of ideas per category), and (e) the ARC. See Figure 2 for an explanation of these variables.

Hypotheses. When group members wait for their turns, delays arise between the generation and articulation of ideas. We suggest that these delays interfere with idea generation in two qualitatively different ways, corresponding to the two stages of idea generation (Figure 3). First, when an idea has been generated, it must be temporarily stored in WM. According to our short-term forgetting assumption, when ideas cannot be expressed immediately they can be forgotten. The longer an idea has to be remembered, the more likely it becomes that it is forgotten. How long ideas need to be remembered is dependent on when ideas are generated (i.e., shortly after the onset of delays or later) and how long the delay is. We predict that relatively long delays will interfere especially with ideas within a train of thought. Because, as we have shown, category repetitions (ideas within a train of thought) are relatively quick, these ideas will have to be remembered longer and thus are more likely to be forgotten during a delay. This will lead to fewer category repetitions and thus shorter trains of thought when delays are relatively long (Figure 3, top panel), resulting in shorter clusters of semantically related ideas (H8). This will also be associated with lower levels of semantic clustering: When there are fewer category repetitions (everything else equal), there will be less semantic organization (H9). Shorter clusters within a semantic category further imply that fewer ideas are generated in that category, and thus lower within-category fluency (on average, fewer ideas per category, H10). Eventually, the consequence is fewer ideas (lower productivity, H11). The matrix model would also predict fewer ideas with longer delays because it also assumes that generated ideas are forgotten during delays (Brown et al., 1998). However, it does not predict that longer delays interfere especially with category repetitions because it does not predict that category repetitions are faster than changes.

The second way in which delays can interfere with idea generation is presented in the bottom panel of Figure 3. The activation of images is a controlled process (requires WM capacity), and a concurrent load on WM will thus interfere with the activation of images (see e.g., Baddeley, 1996). According to our delay-monitoring assumption, delays must be monitored, which takes WM capacity. We hypothesize that the less predictable delays are, the more they need to be monitored, and the more WM capacity is required. Thus, when delays are unpredictable either because it is unpredictable when they begin (e.g., one does not know when fellow group members will express their ideas) or because it is unpredictable when they end (e.g., some ideas take longer to express than others), the activation of images will be disrupted. This will lead not to shorter trains of thought, as with long delays, but to fewer trains of thought. This can be observed as a reduction of the number of semantic clusters (as opposed to shorter clusters) and consequently in fewer category changes (H12). Further, the reduction in the number of clusters may be reflected in a reduction in the number of categories that are surveyed during idea generation (lower idea diversity). However, it may also mean that a person returns less often to a previously used category, in which case the average number of ideas per category (within-category fluency) will be lower. In ei-

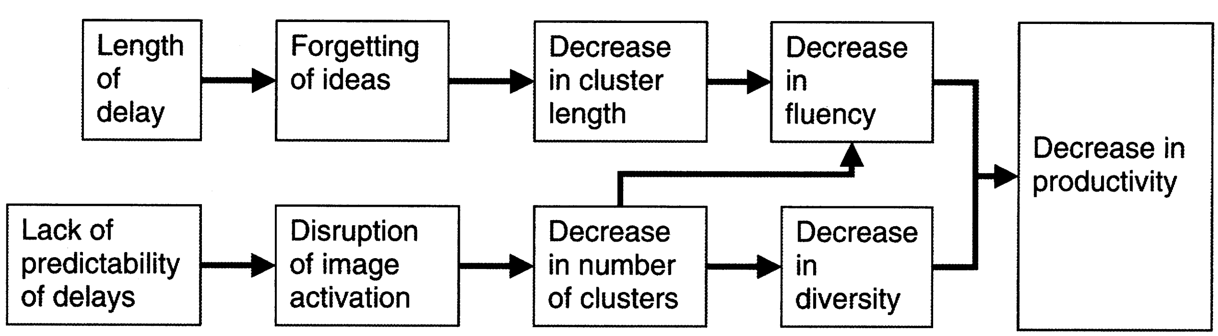

Figure 3. **The two-stage theory of production blocking (from Nijstad et al., 2003).**

Table 3. *Productivity, Cluster Length, Number of Clusters, Diversity, Within-Category Fluency, and Clustering (ARC) as a Function of Delay Length (Adapted from Nijstad, Stroebe, & Lodewijkx, 2003, Exp. 1)*

	Delay Length					
Dependent Variable	**Control (n = 10)**	**1 Sec (n = 10)**	**3 Sec (n = 10)**	**5 Sec (n = 10)**	**7 Sec (n = 10)**	***t*-value Linear Contrast**
Productivity	51.68[a]	49.96[ab]	45.41[ab]	45.39[ab]	37.96[b]	2.20*
	(14.63)	(12.60)	(11.84)	(18.10)	(14.58)	
Cluster Length	1.59[a]	1.34[b]	1.36[b]	1.35[b]	1.29[b]	2.39*
	(.38)	(.23)	(.11)	(.21)	(.19)	
Number of Clusters	32.56[a]	38.99[a]	34.47[a]	36.10[a]	29.89[a]	0.80
	(8.37)	(10.27)	(9.99)	(15.56)	(9.73)	
Diversity	15.15[ab]	17.57[a]	16.34[ab]	15.37[ab]	14.67[b]	1.13
	(2.69)	(2.26)	(3.23)	(2.95)	(2.83)	
Within-Category Fluency	3.38[a]	2.86[ab]	2.75[b]	2.93[ab]	2.60[b]	2.09*
	(.65)	(.78)	(.51)	(.72)	(.72)	
ARC	.41[a]	.28[ab]	.33[ab]	.31[ab]	.23[b]	2.41*
	(.18)	(.16)	(.10)	(.13)	(.11)	

Note: Standard deviations are in parentheses. Different superscripts indicate a significant difference on a post hoc test (LSD).
*$p < .05$.

ther case, productivity should be lower as compared with a condition without delays (H13). Any factor (besides delay predictability) that causes a concurrent load of WM in fact should have similar effects, including the factors manipulated by Diehl and Stroebe (1991; e.g., a combination of a speakers' list and communication).[7] Note that the matrix model does not predict this because it does not assume active search processes involving WM.

Results. These hypotheses were tested in several studies (Nijstad et al., 2003). In these studies, we had individual participants generate ideas at a computer terminal and confronted some of them with electronic delays. Using computers allowed us to have precise control over the length and predictability of delays. Individuals generated ideas on the environment problem in a 20-min session, following Osborn's (1957) rules. Some of them (control group) could enter their ideas whenever they wanted. Others were confronted with delays, during which the screen was blank and they could not enter ideas. No individual had access to ideas of others during delays; they simply had to wait in front of a blank screen.

In a first study (Nijstad et al., 2003, Experiment 1), we used very predictable delays. Participants were confronted with one delay each time they wanted to enter an idea, and for each participant delay length was always the same. Thus, each participants had to wait in front of a blank screen for a fixed length of time, after which a screen appeared in which they could enter one idea. Delay length was manipulated as a between-participants factor. Some were confronted with relatively short delays, and others with longer delays. Delay length was set at 1, 3, 5, or 7 sec (dependent on condition) because previous research had shown that a delay of 5 sec was sufficient to obtain a blocking effect in electronic brainstorming (Gallupe et al., 1994). Finally, there was a control group that was not confronted with delays. All ideas were reliably categorized, and dependent variables were productivity, cluster length, number of clusters, diversity, within-category fluency, and clustering (ARC; see Figure 2 for these measures).

Results were consistent with predictions. Longer delays led to shorter clusters, fewer ideas per category, lower clustering (ARC), and fewer ideas (see Table 3), confirming H8 through H11. However, the (very predictable) delays had no effect on the number of clusters or on diversity (the number of categories surveyed), presumably because little cognitive capacity was required to monitor delays. According to our model (see Figure 3), the effects of delay length on within-category fluency and productivity should be mediated by cluster length. We therefore also performed mediation analyses using the regression approach suggested by Baron and Kenny (1986). We indeed found evidence for mediation: Delay length had a direct effect on the dependent variables, productivity ($\beta = -.28$, $p < .05$) and within-category fluency ($\beta = -.26$, $p < .05$), and had a direct effect on cluster length (mediator, $\beta = -.32$, $p < .05$). When controlling for cluster length, the direct effects of delay length on productivity ($\beta = -.15$) and fluency ($\beta = -.10$) ceased to be significant, whereas the effect of cluster length was significant ($\beta = .49$, $p < .001$). Thus, delay length had its effects on productivity through a reduction of cluster length. This

[7]A similar effect has been obtained in free recall of categorized lists (Cinan, 2003). Cinan found that a concurrent load on WM during retrieval (through dual task performance) led to fewer clusters and fewer category changes, presumably because it interfered with the cue-dependent activation of material in LTM. However, it had no effect on cluster length or on category repetitions, presumably because the recall of items within categories is relatively effortless and automatic and requires less WM capacity.

first study thus suggests that relatively long delays interfere with a person's train of thought, presumably because ideas within a train of thought were forgotten during longer delays, leading to shorter clusters, lower levels of clustering, and fewer ideas.

In another study (Nijstad et al., 2003, Experiment 3), we compared unpredictable and predictable delays. There were three conditions: (a) a no delay control condition, (b) a condition with fixed delays, and (c) a condition with random delays. In the no delay control condition, participants were asked to generate as many ideas as possible and were not confronted with delays. In the fixed delay condition, participants were confronted with a 7 sec delay each time they wanted to enter an idea (identical to the 7 sec condition of Experiment 1). In the random delay condition, a total of 60 delays were randomly distributed across the session. Thus, each time the participant wanted to enter an idea, there was a chance that he or she was blocked. Because of this random distribution of delays, it was possible that participants could sometimes enter two or more ideas before they were blocked, and it was also possible that they were blocked several times before they could enter an idea. This made the onset of delays unpredictable. Further, delays could vary randomly in length in that condition (between 2 sec and 12 sec, with an average of 7 sec), which made delay endings unpredictable.

In the fixed delays condition, we replicated the results of Experiment 1 (Table 4). Thus, fixed delays reduced cluster length, and this in turn led to a decrease in productivity. As in Experiment 1, fixed delays had no effect on the number of clusters. In the random delays condition, on the other hand, we found a different pattern. Delays in that condition had no effect on cluster length but did lead to a lower number of clusters and fewer ideas, confirming H12 and H13 (Table 4). Additional analyses showed that random delays had no effect on cluster length because participants were able to maintain their train of thought within a category when they had the opportunity to enter several ideas before they were blocked (which was only possible in the random delay condition). In that condition, we found that the first idea after a delay was less often a category repetition than the second or later idea after a delay. Thus, when participants could enter several ideas before they were blocked, there were more category repetitions, longer clusters, and a higher level of clustering.

As in Experiment 1, we also performed mediation analyses. In the fixed delay condition, we found that cluster length mediated the effects of delays, similar to what we found in Experiment 1. Thus, a dummy variable comparing the fixed delay condition with the control condition had an effect on productivity (dependent variable, $\beta = -.41, p < .05$) and on cluster length (mediator, $\beta = -.51, p < .01$). When cluster length was controlled for in a third regression, the effect of the dummy variable on productivity was no longer significant ($\beta = -.12$), whereas cluster length had a significant effect ($\beta = .57, p < .001$). In the random delays condition, however, the number of clusters mediated the effect of delays on productivity. A dummy variable comparing the random delays condition with the control condition had an effect on productivity (dependent variable, $\beta = -.45, p < .01$) and on the number of clusters (mediator, $\beta = -.48, p < .01$). When we controlled for the number of clusters, the effect of the dummy variable on productivity ceased to be significant ($\beta = -.03$), and the effect of the number of clusters was significant ($\beta = .87, p < .001$). Thus, consistent with SIAM, two different processes explain the effects of delays: (a) a reduction of cluster length when delays are relatively long, and (b) a reduction in the number of clusters when delays are unpredictable.

Conclusion. These results supported our predictions that production blocking interferes with idea gen-

Table 4. *Productivity, Cluster Length, Number of Clusters, Diversity, Within-Category Fluency, and Clustering (ARC) as a Function of Delays (Adapted from Nijstad, Stroebe, & Lodewijkx, 2003, Exp. 3)*

	Delay Condition			
Dependent Variable	Control	Fixed Delays	Random Delays	F-Value ANOVA
Productivity	45.78[a]	35.76[b]	35.67[b]	6.67**
	(11.95)	(10.46)	(9.82)	
Cluster Length	1.36[a]	1.22[b]	1.41[a]	2.86†
	(.12)	(.12)	(.32)	
Number of Clusters	33.26[a]	29.32[ab]	26.22[b]	5.52**
	(6.84)	(8.09)	(8.12)	
Diversity	15.24[a]	16.12[a]	16.24[a]	0.18
	(3.26)	(3.90)	(3.95)	
Within-Category Fluency	3.01[a]	2.20[b]	2.21[b]	16.85***
	(3.03)	(2.21)	(2.18)	
ARC	.29[ab]	.22[b]	.36[a]	2.93†
	(.10)	(.15)	(.23)	

Note: Standard deviations are in parentheses. Different superscripts indicate a significant difference on a post hoc test (LSD).
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

eration in both stages of the process. Whereas the matrix model could potentially explain the effects of delay length on productivity, it would have trouble accounting for the effects of delay predictability. SIAM, on the other hand, predicts both sets of effects because it assumes that knowledge activation involves the central executive component of WM. Unpredictable delays seem to interfere with this controlled activation of knowledge in LTM.

It should be noted that we have obtained these results under rather artificial conditions: Individuals worked at computers and were confronted with electronic delays. This raises the question whether the same processes operate in real groups. Several findings suggest that they do. First, we have also run a study in which we told half of our participants (all working individually at computers) that they blocked each other, which in fact was not true (Nijstad, 2000). Thus, some participants believed that each time they were blocked it was because another group member was typing at that time. This manipulation had no effects on performance. Second, Diehl (1991) found that the level of clustering (ARC) of group members in real groups was markedly lower than that of individuals, providing some support for the first mechanism we propose. We would suggest that the blocking effect even is exacerbated in real groups, because in real groups more sources of distraction (such as communication) are present than in our paradigm (cf. Diehl & Stroebe, 1991). According to our reasoning, this should further interfere with idea production because it interferes with the controlled activation of knowledge in LTM. However, this second mechanism has not yet been tested in real groups.

Idea Exposure: The Cognitive Stimulation Hypothesis

Background. In our blocking studies, participants had no access to ideas of others. As we discussed earlier, however, evidence suggests that when production blocking is eliminated, as in electronic brainstorming and brainwriting, reading the ideas of others may in fact be stimulating and lead to productivity gains (Dennis & Valacich, 1993; Dugosh et al., 2000; Paulus & Yang, 2000). Further, Dugosh et al. have obtained stimulation effects in an exposure paradigm: Individual participants who listened to an audiotape containing stimulus ideas while they were generating ideas were more productive. In this section, we develop and test hypotheses about why this is the case.

Hypotheses. Our external cue assumption states that ideas of others, provided that they are attended to, are added to the search cue to activate images in LTM. We agree with Diehl et al. (2002) that this should have effects similar to those of experimenter-provided cues in the part-list cueing paradigm. Thus, ideas of others will be helpful if they provide access to idea categories that otherwise would not be accessible. However, we do not agree with Diehl et al. that presenting idea exemplars should invariably lead to lower levels of within-category fluency.

There are two interrelated reasons why idea exchange in groups will not have detrimental effects on within-category fluency in an idea generation task. First, ideas are exchanged continuously throughout the discussion, and not only given as cues at the start of the session. Second, in free recall there usually are a small number of items to remember in each of the categories, but in idea generation there are no a priori limitations on the number of ideas per category. If the same categories are reactivated throughout the session, because stimulation ideas from these categories are offered continuously, they will lead to high numbers of ideas in that category (particularly if there are many potential ideas in that category). In fact, so many ideas will be generated in the stimulated categories that overall within-category fluency (the average number of ideas across all surveyed categories) will be enhanced. This is consistent with findings in the part-list cueing paradigm, in which retrieval of words that are strongly associated to the retrieval cues is enhanced—it is the retrieval of other words that is harmed (e.g., Roediger, 1978).

Ideas of others should aid idea generation in two different ways. First, if ideas of others help retrieve categories of ideas, they will increase the diversity of idea production (more categories are surveyed). This will happen when the stimulation ideas are diverse (from a wide range of semantic categories), as Diehl et al. (2002) have found (H14). Second, if ideas are homogeneous (from only a few categories), they will lead to a high number of ideas in those few categories, and thus to high within-category fluency (H15). This will, however, be at the expense of within-category fluency in the remaining categories or at the expense of fewer categories being surveyed (H16).

The matrix model would also make these predictions. According to that model, stimulation ideas increase activation of ideas in stimulated categories. This should also lead to more categories being surveyed with semantically diverse stimulus ideas, and many ideas being generated within the stimulated categories with homogeneous stimulus ideas. However, SIAM can make an additional prediction regarding response latencies that the matrix model would not make. In the matrix model, stimulation ideas directly activate semantically related ideas. SIAM, on the other hand, assumes that stimulus ideas are added to the search cue. Stimulus ideas should thus reduce the time needed to develop these search cues because the elements of the search cue do not have to be generated but are readily

Table 5. *Productivity, Diversity, Within-Category Fluency, and Response Latencies by Condition (Adapted from Nijstad, Stroebe, & Lodewijkx, 2002)*

Measure	No Stimulation Control ($N = 15$)	Homogeneous Stimulation ($N = 24$)	Diverse Stimulation ($N = 24$)	F-Value ANOVA
Productivity	32.40[a] (10.54)	39.83[b] (10.06)	40.29[b] (10.54)	3.14*
Diversity	14.27[a] (2.84)	13.04[a] (2.44)	17.67[b] (3.82)	13.85***
Within-Category Fluency	2.24[a] (.42)	3.08[b] (.63)	2.28[a] (.35)	20.67***
Within Category Fluency Corrected	2.24[b] (.42)	1.74[a] (.55)	2.28[b] (.35)	10.26***
Response Latency Category Change (sec)	41.53[a] (12.06)	23.86[b] (3.94)	24.94[b] (6.70)	27.49***
Response Latency Category Repetition (sec)	29.33[a] (8.78)	22.78[a] (6.14)	26.21[a] (9.85)	2.69†

Note: Standard deviations are in parentheses. Within-category fluency corrected is the average number of ideas per category when in the homogeneous stimulation condition the two stimulated categories were excluded. Response latencies are in sec. Different superscripts indicate significant differences on a LSD post hoc test.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

available. Previously we argued that category changes are slower than category repetitions because a search cue must be generated before a category change. Stimulus ideas should thus reduce the response latencies of category changes, whereas they should not affect response latencies of category repetitions (H17).

Results. We performed an experiment to test these predictions (Nijstad et al., 2002), in which individuals worked at computers. We used an idea exposure paradigm, and participants were continuously provided with stimulation ideas (cf. Dugosh et al., 2000). We manipulated the content of stimulus ideas and created three conditions: (a) a no stimulus control condition, (b) a condition with diverse stimuli, and (c) a condition with homogeneous stimuli.[8] In the control condition, participants simply generated as many ideas as possible. In the experimental conditions, a stimulus idea was displayed on the screen each time the participant entered an idea. These stimuli were selected from the ideas that other participants had generated earlier and were randomly drawn from a prepared file (but see Footnote 8). In the homogeneous condition, stimuli came from two semantic categories. These two categories were sufficiently "rich" and contained more than 50 different ideas in the earlier experiment. In the diverse stimuli condition, ideas came from 34 different categories (a category system with 50 categories was used). Each category contained at least 5 ideas in the previous experiment. Participants were instructed to pay close attention to stimuli because they would later be tested for recall.

Table 5 lists the results. We found evidence for cognitive stimulation: Productivity (the number of ideas generated) was higher in both conditions with idea exposure than in the control condition. Further, as we anticipated (H14), productivity gains in the diverse stimulation condition were due to the fact that more categories were surveyed than in the control or homogeneous stimulation condition. Diverse stimulation thus increased the accessibility of more idea categories and was therefore helpful. Homogeneous stimulation had no effect on diversity. However, it did lead to more ideas per category, and productivity gains in that condition were due to a higher number of ideas per category, confirming H15. However, we found that within-category fluency for the categories that were not stimulated (the average number of ideas in the other categories, not counting those in the two stimulated categories) was lower in the homogeneous stimulation condition than in the diverse stimulation or control condition. Thus, participants in the homogeneous stimulation condition generated many ideas in the stimulated categories, but this was at the expense of ideas generated in other categories, confirming H16.

We also analyzed response latencies, and results were consistent with H17 (Table 5). The performance increase in both the homogeneous and the diverse stimulation conditions was due to the reduction of response latencies of category changes. Category repetitions were about equally fast in the three conditions. How-

[8]There was a second manipulation in that study, which was a manipulation of stimulus sequence. Some participants were confronted with a random sequence of stimulus ideas. However, others were confronted with idea clusters, in which five successive ideas were from the same category (clusters were drawn at random). This manipulation only affected the ARC and is not discussed further (for details, see Nijstad et al., 2002).

ever, category changes in the stimulation conditions were as fast as category repetitions, and faster than category changes in the control condition. According to us, this implies that participants in the stimulation conditions needed less time to activate images in a new category. We suggest that the reason is that they did not have to generate their own search cues because stimulation ideas were readily available and could be added to the search cue. This aided the activation of images in a new domain, that is, before a category change.

Conclusion. These findings, and findings reported by others (e.g., Dugosh et al., 2000), clearly suggest that cognitive stimulation is possible. Stimulation ideas can increase the number of categories surveyed and thus may help to activate categories of ideas. Further, continuous stimulation of relatively rich categories leads to more ideas being generated in those categories, but this is at the expense of fewer ideas being generated in other categories. Our analysis of response latencies shows that stimulation is effective because it helps participants to make quick category changes, which is consistent with the idea that stimulus ideas are added to the search cue and aid the activation of problem-relevant knowledge. The matrix model would have more difficulty in accounting for that last finding because it does not use search cues. In sum, the evidence is consistent with what SIAM would predict.

As was the case in our blocking studies, findings have been obtained under artificial conditions in which cues were presented by the experimenter rather than by fellow group members. Again, this raises the question whether results generalize to real groups. Previously we have discussed evidence for cognitive stimulation and productivity gains in groups with idea sharing (EBS and brainwriting groups), suggesting that stimulation effects are not restricted to experimenter-provided cues. Unfortunately, in none of these studies have ideas been categorized, preventing conclusions on whether the gains were due to increased idea diversity or increased within-category fluency.

Ideas have been categorized in several other studies. Larey and Paulus (1999) have categorized ideas generated by interactive and nominal groups. They found that interactive groups generated fewer ideas and surveyed fewer categories than did nominal groups. However, that "cognitive uniformity" effect (Lamm & Trommsdorff, 1973) may either have been caused by production blocking, or by overhearing ideas of others. Ziegler, Diehl, and Zijlstra (2000) have categorized the ideas generated by electronic groups without production blocking. They also found that groups that could share ideas surveyed fewer categories than nominal groups in which idea sharing was not possible. Productivity did not differ between these conditions, suggesting that the groups that did share ideas surveyed their categories in greater depth (higher within-category fluency).

The cognitive uniformity effect may be restricted to relatively homogeneous groups. Diehl (1991; also see Stroebe & Diehl, 1994) manipulated group heterogeneity in interactive and nominal groups. Based on dominant associations of participants to a certain topic, Diehl created homogeneous (much overlap in associations) and heterogeneous groups (little overlap in associations). He found that only homogeneous interactive groups surveyed fewer categories of ideas, but heterogeneous interactive groups surveyed as many categories as their nominal counterparts. This suggests that the cognitive uniformity effect may be limited to relatively homogeneous groups. It further suggests that heterogeneous groups may outperform nominal groups if ideas are not shared verbally but through an EBS system or through written notes. Valacich, Wheeler, Mennecke, and Wachter (1995) manipulated group heterogeneity in EBS groups and found that heterogeneous groups outperformed homogeneous groups in large groups (eight members or more). Unfortunately, they did not have a control group without idea sharing, so no firm conclusions are possible as yet.

Cognitive Failures: The Reduction of Failures Hypothesis

Background. As discussed earlier, although (verbally) interactive groups produce fewer ideas than nominal groups, group members usually are more satisfied with their performance than individuals. Further, most people are convinced that group brainstorming is more effective than individual brainstorming, a phenomenon that has been labeled *the illusion of group productivity* (Paulus et al., 1993; Stroebe et al., 1992). The two explanations that have been offered for these findings are memory confusion (Stroebe et al., 1992) and social comparisons (Paulus et al., 1993). We suggest a third explanation of the illusion of group productivity, and we suggest that this explanation also has consequences for task enjoyment and brainstorming persistence of individuals and group members. We hypothesize that group members perceive that idea sharing is stimulating because communication reduces the subjectively experienced number of failures in a group setting. This is one reason why most people believe that group idea generation is effective.

Hypotheses. There are two reasons why idea sharing in groups will lead to a reduction in experienced cognitive failures. First, group members often have to wait for their turns before they can express their ideas. According to our search inertia assumption, during these delays group members will not always start a new search for ideas. Individuals do not share speaking time with others, but instead produce ideas continually throughout a session. This implies that individuals will start more searches than group members, which eventually will lead to a higher number of successful searches (ideas), but also to a higher number of failures (unsuccessful searches). Second, overhearing ideas of others can be stimulating, and group members also report that they found the ideas of others stimulating (Stroebe et al., 1992). We have argued that ideas of others facilitate the activation of images, and we may expect that fewer attempts to activate an image result in a failure (cf. Figure 1). Thus, group members start fewer searches and, due to mutual stimulation, fewer of these searches will result in failures. As a consequence, people experience fewer failures in a group setting than in an individual setting (H18).

This reduction of failures in a group session is predicted to have several effects. First, as argued before, a brainstorming session will not be very enjoyable when many trials result in failures. Thus, group members should enjoy the brainstorming session more than individuals, which in general is true (e.g., Paulus et al., 1993; Nijstad et al., 1999). More important, that effect should be mediated by experienced failures (H19). Second, when many searches result in failures, the impression will arise that one is not doing particularly well. Hence, failures should be associated with satisfaction, and more failures experienced by individuals should explain why they are less satisfied with their performance than group members (despite their higher productivity; H20). Finally, throughout the session the number of searches resulting in failures will rise. This will lead to the expectation that not many more ideas are possible (low expectancy) and that the session might as well be ended. If group members experience fewer failures than individuals, groups should reach the point at which they want to end the session later than individuals do, and should therefore be more persistent (H21).

In our studies, we have used groups of different sizes to see how group size affects failures, satisfaction, enjoyment, and persistence. It is important to distinguish between the effects of group size on individual level variables (such as failures, enjoyment, and satisfaction) and group level variables (persistence). At the individual level, it is difficult to generate specific hypotheses regarding the effects of group size on failures (and thus on enjoyment and satisfaction). On the one hand, in larger groups the blocking effect is exacerbated (more people have to share speaking time), which would lead to even fewer searches being started, and thus fewer failures. On the other hand, more ideas are mentioned in larger groups (although not per person). This increases the chance that an idea is generated that another group member has already mentioned, which also is a failure. Thus, the relation between group size and experienced failures is difficult to predict beforehand. However, for persistence at the group level, we expect a positive relation with group size. The larger the group, the greater the chance that at least one group member is still able to generate ideas. As long as at least one person is still productive, one can expect the group to continue their session. Thus, larger groups will be more persistent than smaller groups (H22).

Results. Our predictions have been tested in several studies (Nijstad, Stroebe, & Lodewijkx, 1999, 2006). In these studies, participants were instructed to continue brainstorming until they felt it was a good time to stop. In one experiment, they brainstormed alone or in a dyad, 4-person, or 6-person group. In all conditions (including the alone condition), participants articulated their ideas aloud, and all sessions were recorded on audiotape. We measured participants' persistence (how many minutes they spent at brainstorming) and productivity (number of ideas; coded from the tapes). In addition, we measured cognitive failures, satisfaction, and enjoyment in a postexperimental questionnaire (see a previous section for more information).

Consistent with predictions (H21 and H22), we found that groups were more persistent than individuals and that persistence increased with group size: Our individuals on average brainstormed for 30 min, but our 6-person groups continued for about 45 min (Table 6). Groups were more productive than individuals, and productivity increased with group size because more people contributed ideas in larger groups. However, we also created nominal groups by pooling the ideas of randomly chosen individuals. When the production of equivalent-size nominal groups was subtracted from the production of the real groups, we found that interactive groups still were not as productive as nominal groups, and productivity losses increased with group size. Apparently, even when persistence is taken into account, the blocking effect still caused productivity losses in relatively large groups.

Further, and consistent with H18, we found that group interaction reduced the number of experienced failures: Individuals indicated that they had experienced more failures than did group members, but members of groups of different sizes did not differ. Further, we found that group members, irrespective of group size, were more satisfied with their performance and enjoyed the brainstorming more than individuals (Table 6). Finally, mediation analyses showed that failures mediated the effect of type of setting (group vs. individual) on both satisfaction and enjoyment. Thus, type of setting had effects on satisfaction ($\beta = .31, p < .001$), enjoyment ($\beta = .25, p < .01$), and failures (the mediator, $\beta = -.37, p < .001$). When statistically controlling for failures, the effect of setting on satisfaction was reduced ($\beta = .17$), and the effect of failures was significant ($\beta = -.37, p < .001$). Similarly, the effect of

Table 6. *Persistence, Productivity, Productivity Loss, Failures, Enjoyment, and Satisfaction as a Function of Group Size (Adapted From Nijstad, Stroebe, & Lodewijkx, 1999, 2006)*

	Group Size			
Measure	**Individuals**	**Dyads**	**Four-Person Groups**	**Six-Person Groups**
Persistence	30.20 (18.32)	34.48 (14.30)	39.76 (12.07)	44.14 (6.48)
Productivity	55.85[a] (40.42)	107.89[b] (43.45)	171.44[c] (77.20)	212.50[c] (88.47)
Productivity Loss	—	–0.43[a] (43.45)	–38.74[ab] (77.20)	–92.67[b] (88.47)
Failures	6.64[a] (1.00)	5.69[b] (1.40)	5.14[b] (1.52)	5.42[b] (1.35)
Enjoyment	6.50[a] (1.45)	7.33[b] (0.75)	7.36[b] (0.67)	7.28[b] (0.41)
Satisfaction	5.15[a] (2.33)	7.17[b] (0.75)	6.47[b] (1.19)	6.50[b] (0.71)

Note: Standard deviations are in parentheses. Persistence in min; productivity and productivity losses in number of ideas. Questionnaire items were measured on 9-point scales. Different superscripts indicate significant differences on a post hoc test.

setting on enjoyment was reduced ($\beta = .15$) when controlling for the effect of failures ($\beta = -.27$, $p < .01$). This confirms H19 and H20.

In a second experiment, we replicated and extended these findings. We again manipulated group size (individuals, dyads, and 4-person groups) but also used two different topics. Topic was manipulated because previous research had shown that one topic was easier than the other. We reasoned that the difficult topic would lead to more failures, and therefore to lower persistence, satisfaction, and enjoyment (H23). We replicated findings of our first study and found that persistence increased with group size. We also found that group interaction reduced the number of failures (individuals reported more failures than members of dyads and four-person groups). Further, and consistent with earlier findings, group members (irrespective of group size) enjoyed themselves more and were more satisfied with their performance than were individuals. As in our other study, failures mediated the effects of type of setting (interactive vs. nominal) on both satisfaction and enjoyment. Thus, type of setting had significant effects on satisfaction ($\beta = .29$, $p < .01$), enjoyment ($\beta = .30$, $p < .001$), and failures ($\beta = -.29$, $p < .01$). When controlling for failures, the effect of setting on satisfaction was reduced to nonsignificance ($\beta = .14$), whereas the effect of failures was significant ($\beta = -.49$, $p < .001$). Similarly, the effect of setting on enjoyment was reduced (but remained significant, $\beta = .22$, $p < .05$) when controlling for failures, which had a significant effect ($\beta = -.34$, $p < .001$). Thus, this study again confirmed H18 through H22.

With regard to the effects of our second manipulation (brainstorming topic), results were also largely consistent with predictions (H23). Participants generated more ideas on the easier topic than on the difficult topic and also indicated that the easy topic was easier than the difficult topic. As expected, participants experienced more failures with the difficult topic than with the easier topic. Further, participants brainstorming on the easy topic enjoyed themselves more and were more satisfied than those brainstorming on the difficult topic. The effects of topic on enjoyment and satisfaction were mediated by failures. Thus, topic had significant effects on enjoyment ($\beta = .29$, $p < .01$), satisfaction ($\beta = .20$, $p < .05$), and failures ($\beta = -.21$, $p < .05$). The effect of topic on enjoyment was reduced but remained significant ($\beta = .20$, $p < .05$) when controlling for failures, which had a significant effect ($\beta = -.34$, $p < .001$). The effect of topic on satisfaction was reduced to nonsignificance ($\beta = .10$) when controlling for failures, which had a significant effect ($\beta = -.49$, $p < .001$).

However, there were no differences in persistence for the two topics. One reason may be that differences in failures between the two topics mainly occurred in the early stages of the session, and later on, when people considered ending the session, the differences had largely leveled out. Indeed, Diehl and Stroebe (1991) found that the productivity differences between nominal and interactive groups mainly occur in the early stages of a session, and the same may be true for differences in topics.

Conclusion. These studies show that group interaction leads to a reduction in the number of failures people experience, and that this reduction of failures is accompanied by higher satisfaction, enjoyment, and persistence. Further, because we were able to manipulate failures through topic difficulty, it seems that failures do play a causal role in this process. It seems likely that this also contributes to the illusion of group productivity (most people erroneously believe group brainstorming to be more effective than individual brainstorming). Because people experience fewer failures when they work in a group, they will conclude that brainstorming is much easier when working with a group (fewer failures) and attribute this to the stimulating value of others' ideas. However, they seem unaware that in a group brainstorming session they often have to wait for their turns, and that during these delays they are unproductive. We should note that there is clear evidence that other processes such as social comparisons also contribute to this illusion (Paulus et al., 1992; also see Nijstad et al., in press)

It also appears that group members keep each other going, resulting in higher levels of persistence and an increase in persistence with group size. One can look at this finding in two ways. First, one may argue that this is a benefit of working in a group, and that groups may not be as unproductive as generally assumed. Indeed, our groups compensated part of their productivity loss by being more persistent. Another way to look at the result is that groups indeed are ineffective: They need more time and still are outperformed by individuals (at least in larger groups). Nevertheless, there are benefits associated with group brainstorming: People enjoy working in groups and are generally quite satisfied with their performance.

General Discussion

Research on group idea generation has revealed some consistent and robust findings. Individuals working alone generate more ideas, and more good ideas, than group members, and this productivity loss of groups is caused by the fact that group members must take turns to express their ideas (production blocking). Yet, people hold the erroneous belief that group brainstorming is more effective than individual brainstorming. However, when turn taking is not necessary (when ideas are shared on written notes or through computers) there is no productivity loss, and even productivity gains can occur. Using a social–cognitive approach, we developed our SIAM model to explain and integrate these findings. We proposed that effects of group interaction can be interpreted as cognitive stimulation and interference of the individual group member's thought process.

SIAM assumes that idea generation is a two-stage process, in which an effortful, cue-dependent stage of knowledge activation is followed by a stage of idea production within a semantic domain (a train of thought). At the individual level, SIAM predicts that (a) there is above chance semantic clustering of ideas, and (b) the generation of ideas within clusters (category repetitions) is faster than the generation of ideas between clusters (category changes), and clustering is positively correlated with productivity (the number of ideas). These predictions were supported in four different studies. Further, the two stages are represented as negative feedback loops, and the search for ideas is ended when the number of cognitive failures (instances in which no new idea is generated) exceeds a stop criterion. We argued that failures should be associated with task enjoyment, satisfaction with performance, and expectancy (the number of ideas one still expects to be able to generate). In six different studies, we found support for these predictions.

At the group level, SIAM can be used to generate predictions on how individual level cognitive processes are affected by idea sharing in groups. First, using some additional assumptions, SIAM suggests that the productivity loss caused by production blocking is due to cognitive interference in both stages of the idea generation process. When delays between the generation and expression of ideas, which arise when someone else is talking, are long, generated ideas may be forgotten, leading to shorter trains of thought (shorter semantic clusters of ideas), lower levels of clustering, and fewer ideas. When delays are unpredictable, they need to be monitored, which consumes cognitive resources. This interferes with the controlled process of knowledge activation and leads to fewer trains of thought and fewer ideas. These predictions were supported in several studies (see Nijstad et al., 2003).

Second, SIAM suggests that the ideas of others, when they are attended to, will be added to a search cue to probe memory. Stimulus ideas should have effects similar to those of experimenter-provided cues in the part-list cueing paradigm (e.g., Slamecka, 1968): They make related knowledge more accessible, whereas knowledge unrelated to the stimulus ideas becomes less accessible. We found support for these predictions: Exposure to ideas makes it more likely that related ideas are generated (Nijstad et al., 2002). Idea sharing may lead to productivity gains, either because inaccessible knowledge becomes accessible or because knowledge remains highly accessible throughout the session. At the same time ideas that are not strongly associated with stimulus ideas are less likely to be generated, and this may sometimes lead to productivity losses: Fewer ideas are being generated in the less accessible semantic categories (also see Diehl et al., 2002; Ziegler et al., 2000). The positive effects of idea sharing are due to a reduction in response latencies of category changes, presumably because stimulation ideas reduce the time needed to develop search cues.

Third, we have argued that failures (instances in which a search for ideas is unsuccessful) are less numerous as a consequence of idea sharing in groups. In groups, speaking time is shared among group members and they therefore start fewer searches than individuals, eventually leading to fewer ideas but also to fewer failures. This reduction of failures in a group setting was found to have several consequences, such as higher levels of satisfaction and enjoyment. Further, many failures occurring toward the end of the session signals that not many new ideas are possible, and people as a consequence want to end the session. Because group members experience fewer failures than individuals, they go on longer, reducing their productivity loss. Because failures are less numerous in a group setting, people come to believe that idea sharing stimulates creativity: Idea generation appears to be easier in a group setting than when brainstorming alone.

Throughout the article, we have contrasted our predictions with those generated by the Brown et al.

(1998) matrix model. The main difference between the two theories is that search processes in SIAM are cue dependent and effortful (they require WM capacity), whereas the matrix model assumes that ideas are activated through spreading of activation. We have argued that the matrix model can account for some of our findings but would have difficulty accounting for others. For example, the matrix model cannot account for the effects of predictability of delays. If no WM capacity is needed to activate ideas, there is no reason to expect that distraction will interfere with idea generation. Further, the matrix model does not necessarily predict category repetitions to be faster than category changes, and therefore does not predict that stimulation reduces the response latencies of category changes but not of category repetitions. We would therefore argue that, although the matrix model has some promising features, SIAM is a more general model of idea generation: It can account for more empirical findings.

In sum, SIAM generates detailed and testable hypotheses. To test these hypotheses, we have used dependent variables that are not routinely assessed in idea generation research, such as idea diversity, cluster length, and the ARC. We argue that this has led to a more detailed understanding of group and individual idea generation. SIAM can account for many research findings, and it appeared that different phenomena in idea generation have similar cognitive underpinnings. Thus, our approach has two benefits: (a) a more detailed understanding of processes underlying idea generation in groups, and (b) an integration of various findings using one explanatory mechanism.

Group Idea Generation Versus Group Recall

There are strong parallels between idea generation and free recall. For example, in both idea generation and free recall there is a decline in performance over time, and there is semantic as well as temporal clustering. Further, group recall is inferior to pooled individual recall (collaborative inhibition; Basden et al., 1997; Weldon & Bellinger, 1997), similar to the productivity loss in brainstorming groups. Clustering in group recall is much lower than in individual recall (Basden et al., 1997), which is also true for group brainstorming (Diehl, 1991). This suggests that common mechanisms might underlie these effects.

Basden et al. (1997) argued that collaborative inhibition is due to a disruption of individual retrieval strategies. Group members who overhear each other abandon their favored order of recall and adopt a less effective one. However, because in the Basden et al. studies group members also had to take turns when recalling items, the effect might also have been caused by production blocking. In idea generation, Diehl and Stroebe (1987, 1991) have shown that overhearing others is not necessary to obtain a blocking effect: Turn taking leads to productivity losses even when participants do not hear each other. If this also is the case in group recall, one would expect turn taking in free recall to have effects similar to those we found for idea generation: shorter clusters with longer delays (due to short-term forgetting) and fewer clusters with unpredictable delays (due to the load on WM; cf. Cinan, 2003). This can clearly be tested in future research.

Some findings obtained by Basden et al. (1997) seem inconsistent with this interpretation. For example, when they had group members recall items category by category, they did not find the collaborative inhibition effect. However, group members still had to take turns recalling items, and there should have been a productivity loss caused by blocking. A procedural difference might be responsible for this. In brainstorming studies, individuals and groups are usually given the same fixed time limit to generate ideas. However, in the Basden et al. studies, participants were not given a time limit, but the session was ended after 30 sec had elapsed in which no item was recalled (see Footnote 2). In our persistence studies, we were able to show that in the absence of fixed time limits groups continue longer than individuals and thereby compensate (part of) their productivity loss. Thus, had a fixed time limit been used in the Basden et al. studies, they might have obtained the collaborative inhibition effect even when group members recalled items category by category.

Nevertheless, the effects of overhearing others may be more negative in free recall than in idea generation. The reason is that the number of items that can be recalled by definition is limited to the learned list, whereas there are no a priori limitations in the number of ideas that can be generated. Any item recalled by a group member necessarily reduces the number of items others can recall and increases the chance that another group member recalls the same item. In idea generation this is true to a lesser extent, and this might be the reason why stimulation rather than inhibition effects of exposure to others' ideas are found. To our knowledge, no such stimulation effects have ever been found for group free recall.

Future Directions

We believe that in future research it is important to explore the dynamic properties of idea sharing in groups. People within groups influence each other, and that process of mutual influence largely determines which ideas are being generated (and how many). It may lead to cognitive uniformity effects (e.g., Ziegler et al., 2000), but this may not always be the case. In particular, the dynamic process of idea sharing in diverse groups deserves attention. When different group members come up with very different ideas, the process of mutual influence is unlikely to result in cognitive uniformity (cf. Diehl, 1991) but may in fact increase the range of ideas generated. Particularly in situations in which groups interact without production blocking, as in electronic brainstorming, idea sharing in diverse groups may have clear benefits for performance.

Second, SIAM assumes that knowledge activation is a controlled and effortful process. Based on this property, one would predict that situations involving a concurrent load on WM (e.g., through dual task performance) would interfere with the activation of knowledge. Because it is presumed that the "central executive" is involved in the process of knowledge activation (Baddeley, 1996; Cinan, 2003), any factor that reduces the available cognitive resources should lead to difficulties when trying to start a new train of thought, and thus to fewer trains of thought. The interfering effects of having to take notes and signing up on a speaking list (Diehl & Stroebe, 1991) might have been caused by the reduced ability to activate knowledge when attention has to be divided among several activities (e.g., taking notes, listening to others, and idea generation). Further, it might be the case that shy people are distracted by irrelevant thoughts (e.g., trying to make a good impression) and therefore have difficulties when brainstorming in a group (see Camacho & Paulus, 1995). These predictions can be tested in future work.

We also think that two specific aspects of how people generate ideas deserve attention. One is how people generate search cues to activate images. We speculated that search cues might be generated through event sampling—thinking about previous experiences. This would imply that events stored in autobiographical memory are retrieved and used to activate images in semantic memory. However, we have no evidence other than informal observations. For example, we had students generate ideas about how the education at the department of psychology might be improved. Many group discussions revolved around group members' specific complaints about the psychology curriculum: They identified problems based on their own experience and then suggested ways to solve these problems. Future research can investigate this, for example by providing people with specific strategies to approach a problem.

A second aspect is how ideas are generated once knowledge has been activated. We think most ideas result from fairly simple operations, such as introduce something that is not there or change something that causes problems. Goldenberg, Mazursky, and Solomon (1999) have suggested a promising approach. They identified a limited number of "creativity templates:" relatively simple operations that result in creative ideas. For example, one creativity template is replacement: removing one component and replacing it with a new one. We think it will be fruitful for creativity research to identify the cognitive operations that lead to creative ideas and look for communalities among these operations.

Brainstorming in Practice

Although the main focus of this article is on theory, we can give a number of practical recommendations. If one wants to stimulate productivity and generate many ideas, it is wise to keep groups with verbal idea sharing small and, if necessary, split up larger groups. Electronic brainstorming and brainwriting are to be preferred above verbal brainstorming. People should be encouraged to pay attention to one another's ideas because these generally are stimulating. People should also be discouraged to engage in off-task discussion or explaining of ideas (cf. Dugosh et al., 2000). Also, short breaks are helpful, and productivity can be enhanced considerably when one distinguishes specific subcategories of a larger problem. Finally, to preserve working memory capacity, external storage of ideas (e.g., a flip-over or notepads) might be helpful. However, as research by Diehl and Stroebe (1987) has shown, taking notes itself may be a source of distraction, and it probably only is beneficial in electronic brainstorming and brainwriting.

Brainstorming assumes that generating a large number of ideas is good, because "quantity breeds quality:" The more ideas are generated, the more good ideas will be found among them. However, what brainstorming practitioners eventually are after is a limited number of high-quality ideas that can subsequently be implemented. Surprisingly little attention has been paid to the stages after brainstorming, such as idea selection and implementation (but see Rietzschel, Nijstad, & Stroebe, 2006). For practice, it is important to know more about how to improve idea selection or how to facilitate implementation. For example, it is not clear whether brainstorming yields real benefits when it comes to eventual idea implementation: Does generating many ideas lead to the eventual implementation of ideas of higher quality? Future research should investigate this.

Conclusion

With our analysis of group idea generation we hope to have shown that the question of "how the group affects the mind" deserves further research efforts. Group idea generation is an area in which theories and findings from cognitive psychology (in particular from memory retrieval) may be fruitfully applied to yield predictions on the effects of intragroup communication on individual level cognitive processes. Indeed, as Hinsz, Tindale, and Vollrath (1997) argued, the utterances of others may both stimulate and interfere with mental processes of group members, which provides "a new area to address the relative impact of process gains (e.g., stimulated cognitive processes…) and process losses (e.g., interference…) in small group task performance" (p. 49). We conclude that integrating theories from cognitive psychology into group research brings valuable insights and may increase our understanding of group level information processing and performance.

References

Amabile, T. M. (1983). Social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology, 45*, 357–376.

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). New York: Freeman.

Baddeley, A. D. (1986). *Working memory.* Oxford, England: Oxford University Press.

Baddeley, A. D. (1990). *Human memory: Theory and practice.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A*, 5–28.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). Oxford, England: Academic.

Baron, R. M., & Kenny, D. A. (1986). The mediator–moderator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Bartis, S., Szymanski, K., & Harkins, S. (1988). Evaluation and performance: A two-edged knife. *Personality and Social Psychology Bulletin, 14*, 242–251.

Basden, B. H., Basden, D. R., Bryner, S., & Thomas, R. L., III. (1997). A comparison of group and individual remembering: Does collaboration disrupt retrieval strategies? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1176–1189.

Boden, M. A. (1990). *The creative mind: Myths & mechanisms.* London: Weidenfeld and Nicolson.

Bousfield, W. A., & Sedgewick, C. H. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology, 30*, 149–165.

Brown, V., Tumeo, M., Larey, T. S., & Paulus, P. B. (1998). Modeling cognitive interactions during group brainstorming. *Small Group Research, 29*, 495–526.

Camacho, L. M., & Paulus, P. B. (1995). The role of social anxiousness in group brainstorming. *Journal of Personality and Social Psychology, 68*, 1071–1080.

Cantor, J., & Engle, R. W. (1993). Working memory capacity as long-term memory activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 5*, 1101–1114.

Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology, 53*, 82–88.

Cinan, S. (2003). Executive processes in free recall of categorized lists. *Learning and Motivation, 34*, 240–261.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Coskun, H., Paulus, P. B., Brown, V., & Sherwood, J. J. (2000). Cognitive stimulation and problem presentation in idea-generating groups. *Group Dynamics, 4*, 307–329.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Learning and Verbal Behavior, 19*, 450–466.

Dennis, A. R., & Valacich, J. S. (1993). Computer brainstorms: More heads are better than one. *Journal of Applied Psychology, 78*, 531–537.

Dennis, A. R., Valacich, J. S., Connolly, T., & Wynne, B. E. (1996). Process structuring in electronic brainstorming. *Information Systems Research, 7*, 268–277.

Diehl, M. (1991). *Kollektive Kreativität: Zur Quantität und Qualität der Ideenproduktion in Kleingruppen.* Unpublished Habilitationschrift, University of Tübingen, Tübingen, Germany.

Diehl, M., Munkes, J., & Ziegler, R. (2002, June). *Brainstorming and cognitive stimulation: When does being exposed to the ideas of others facilitate or inhibit one's own idea generation?* Paper presented at the conference of the European Association of Experimental Social Psychology, San Sebastian, Spain.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology, 53*, 497–509.

Diehl, M., & Stroebe, W. (1991). Productivity loss in idea-generating groups: Tracking down the blocking effect. *Journal of Personality and Social Psychology, 61*, 392–403.

Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yang, H.-C. (2000). Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology, 79*, 722–735.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365–395). Cambridge, MA: MIT Press.

Gallupe, R. B., Bastianutti, L. M., & Cooper, W. H. (1991). Unblocking brainstorms. *Journal of Applied Psychology, 76*, 137–142.

Gallupe, R. B., Cooper, W. H., Grisé, M. L., & Bastianutti, L. M. (1994). Blocking electronic brainstorms. *Journal of Applied Psychology, 79*, 77–86.

Goldenberg, J., Mazursky, D., & Solomon, S. (1999). Toward identifying the inventive templates of new products: A channeled ideation approach. *Journal of Marketing Research, 26*, 200–210.

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 3–28.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 225–240.

Haarmann, H., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin and Review, 8*, 568–578.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin, 121*, 43–64.

Homma, M., Tajima, K., & Hayashi, M. (1995). The effects of misperception of performance in brainstorming groups. *Japanese Journal of Experimental Social Psychology, 34*, 221–231.

Johnson, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kanekar, S., & Rosenbaum, M. E. (1972). Group performance on a multiple-solution task as a function of available time. *Psychonomic Science, 27*, 331–332.

Kerr, N. L., MacCoun, R., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review, 103*, 687–719.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*, 623–655.

Lamm, H., & Trommsdorff, G. (1973). Group versus individual performance on a task requiring ideational proficiency (brainstorming): A review. *European Journal of Social Psychology, 3*, 362–388.

Larey, T. S., & Paulus, P. B. (1999). Group preference and convergent tendencies in small groups: A content analysis of group brainstorming performance. *Creativity Research Journal, 12*, 175–184.

Loftus, E. F. (1974). Activation of semantic memory. *American Journal of Psychology, 86*, 331–337.

Martin, L. L., Ward, D. W., Achee, J. W., & Wyer, R. S. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology, 64*, 317–326.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 69*, 220–232.

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology, 12*, 3–24.

Nagasundaram, M., & Dennis, A. R. (1993). When a group is not a group: The cognitive foundation of group idea generation. *Small Group Research, 24*, 463–489.

Nijstad, B. A. (2000). *How the group affects the mind: Effects of communication in idea generating groups*. Unpublished doctoral dissertation, Utrecht University.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (1999). Persistence of brainstorming groups: How do people know when to stop? *Journal of Experimental Social Psychology, 35*, 165–185.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2002). Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology, 38*, 535–544.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology, 39*, 531–548.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2006). The illusion of group productivity: A reduction of failures explanation. *European Journal of Social Psychology, 36*, 31–48.

Osborn, A. F. (1953). *Applied imagination*. New York: Scribner's.

Osborn, A. F. (1957). *Applied imagination* (2nd ed.). New York: Scribner's.

Parnes, S. J., & Meadow, A. (1959). Effects of "brainstorming" instructions on creative problem solving by trained and untrained subjects. *Journal of Educational Psychology, 50*, 171–176.

Paulus, P. B., Dugosh, K. L., Dzindolet, M. T., Coskun, H., & Putman, V. L. (2002). Social and cognitive influences in group brainstorming: Predicting production gains and losses. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 299–325). Chichester, England: Wiley.

Paulus, P. B., & Dzindolet, M. T. (1993). Social influence processes in group brainstorming. *Journal of Personality and Social Psychology, 64*, 575–586.

Paulus, P. B., Dzindolet, M. T., Poletes, G., & Camacho, L. M. (1993). Perception of performance in group brainstorming: The illusion of productivity. *Personality and Social Psychology Bulletin, 19*, 78–89.

Paulus, P. B., Larey, T. S., Putman, V. L., Leggett, K. L., & Roland, E. J. (1996). Social influence process in computer brainstorming. *Basic and Applied Social Psychology, 18*, 3–14.

Paulus, P. B., & Yang, H. C. (2000). Idea generation in groups: A basis for creativity in organizations. *Organizational Behavior and Human Decision Processes, 82*, 76–87.

Raaijmakers, J. G. W. (1993). The story of the two-store model of memory: Past criticisms, current status, and future directions. In D. E. Myers & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 467–488). Cambridge, MA: MIT Press.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*, 93–134.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2006). Productivity is not enough: A comparison of interactive and nominal groups on idea generation and selection. *Journal of Experimental Social Psychology, 42*, 244–251.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory and Cognition, 2*, 261–269.

Roediger, H. L. (1978). Recall as a self-limiting process. *Memory and Cognition, 6*, 54–63.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76*, 45–48.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192.

Simonton, D. K. (2003). Scientific creativity as stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin, 129*, 475–494.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76*, 504–513.

Smith, S. M. (2003). The constraining effects of initial ideas. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 15–31). New York: Oxford University Press.

Smith, S. M., & Blankenship, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society, 27*, 311–314.

Smith, S. M., & Vela, E. (1991). Incubated reminiscence effects. *Memory & Cognition, 19*, 168–176.

Stasser, G., & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology, 53*, 81–93.

Steiner, I. D. (1972). *Group process and productivity*. New York: Academic.

Stroebe, W., & Diehl, M. (1994). Why groups are less effective than their members: On productivity losses in idea-generating groups. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 5, pp. 271–303). London: Wiley.

Stroebe, W., Diehl, M., & Abakoumkin, G. (1992). The illusion of group effectivity. *Personality and Social Psychology Bulletin, 18*, 643–650.

Sutton, R. I., & Hargadon, A. (1996). Brainstorming groups in context. *Administrative Science Quarterly, 41*, 685–718.

Taylor, D. W., Berry, P. C., & Block, C. H. (1958). Does group participation when brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly, 3*, 23–47.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381–391.

Valacich, J. S., Dennis, A. R., & Connolly, T. (1994). Idea-generation in computer-based groups: A new ending to an old story. *Organizational Behavior and Human Decision Processes, 57*, 448–467.

Valacich, J. S., Wheeler, B. C., Mennecke, B. E., & Wachter, R. (1995). The effects of numerical and logical group size on computer-mediated idea generation. *Organizational Behavior and Human Decision Processes, 62*, 318–329.

Vroom, V. (1964). *Work and motivation*. New York: Wiley.

Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1160–1175.

West, M. A., & Anderson, N. (1996). Innovation in top management teams. *Journal of Applied Psychology, 81*, 680–693.

Ziegler, R., Diehl, M., & Zijlstra, G. (2000). Idea production in nominal and virtual groups: Does computer-mediated communication improve group brainstorming? *Group Processes and Intergroup Relations, 3*, 141–158.